

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2005, 19:45:35 ; Search time 4020 Seconds
(without alignments)
8425.422 Million cell updates/sec

Title: US-10-049-710A-1
Perfect score: 699
Sequence: 1 atggcagtcgcacagacg.....agggcgctccgacagtag 699

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	699	100.0	699	6	BD105336 Method fo
2	699	100.0	1755	1	D3251 Streptomyc
3	267.2	38.2	2234	1	AB001683 Streptom
4	229	32.8	6727	1	SCO007731 Streptom
5	229	32.8	290850	1	AL939127 Streptom
6	213.6	30.6	1328	1	STMAFR Streptomyc
7	211.4	30.2	33064	2	AY338477 Streptom
8	211.4	30.2	210614	1	AB088224 Streptom
9	199.2	28.5	1601	1	AB021882 Streptom
10	197.8	28.3	654	1	AB121071 Kitasatos
11	197.8	28.3	6997	1	AB126048 Kitasatos
12	194.4	28.0	901	1	AY256849 Streptom
13	194.4	27.8	2637	1	AF156161 Streptom
14	180.8	25.9	92294	1	AY117439 Streptom
15	173.4	24.8	1101	1	AB106894 Streptom
16	163	23.3	1129	1	AY026762 Streptom
17	157.8	22.6	104578	1	AY502076 Rhodococc
18	157.4	22.5	300800	1	AP005036 Streptom
19	154.8	22.1	10467	1	AF145049 Streptom

20	145.8	20.9	1340	1	AB000385 Streptomy
21	145.6	20.8	295150	1	SC0939126 Streptomy
22	138	19.7	799	1	AB000384 Streptomy
23	126	18.0	12905	1	AF055922 Streptomy
24	112.4	16.1	300800	1	AP005036 Streptomy
25	93.4	13.4	300550	1	AP005030 Streptomy
26	92.2	13.2	104578	1	AY502076 Rhodococc
27	90.6	13.0	1498	1	AB001609 Streptomy
28	88.6	12.7	1819	6	AX708269 Sequence
29	88.6	12.7	1819	6	AX708269 Sequence
30	85.8	12.3	9521	6	AX188771 Sequence
31	85.8	12.3	9521	6	AX188772 Sequence
32	85.8	12.3	22818	1	SC0276673 Streptomy
33	85.8	12.3	178073	1	SC0590464 Streptomy
34	85.6	12.2	110000	1	AP006618_50 Continuation (51 o
35	82	11.7	3729	1	AB035547 Streptomy
36	79.8	11.4	321250	1	SC0939111 Streptomy
37	78.8	11.3	110000	1	AP006618_06 Continuation (7 of
38	78.8	11.3	110000	1	AP006840_24 Continuation (25 o
39	78.2	11.2	4257	6	AR144708 Sequence
40	78.2	11.2	4257	6	AR145616 Sequence
41	78.2	11.2	4257	6	AR526866 Sequence
42	78.2	11.2	6633	14	HEH5VIG3 X06461 Herpes simp
43	78.2	11.2	26245	14	HS1US L00036 Human herpe
44	78.2	11.2	26245	14	HS1US L00036 Human herpe
45	78.2	11.2	152261	14	HE1CG X14112 Human herpe

ALIGNMENTS

RESULT 1	BD105336	Method for induction of gene expression in plants and plants thereof	699 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD105336	Method for induction of gene expression in plants and plants thereof	699 bp	DNA	linear	PAT 27-AUG-2002
ACCESSION	BD105336	Method for induction of gene expression in plants and plants thereof	699 bp	DNA	linear	PAT 27-AUG-2002
VERSION	BD105336.1	GI:22650910				
KEYWORDS	WO 0196581-A/1.					
SOURCE	Streptomyces virginiae					
ORGANISM	Streptomyces virginiae					
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.					
AUTHORS	Shimmo,A., Kato,K., Yamada,Y., Nihira,T. and Shindo,T.					
TITLE	Method for induction of gene expression in plants and plants thereof					
JOURNAL	Patent: WO 0196581-A 1 20-DEC-2001; KANEKA CORP,ATSUHIKO SHIMMO,KO KATO,YASUHIRO YAMADA,TAKUYA NIHIRA,TAKUYA SHINDO					
COMMENT	OS Streptomyces virginiae					
	PN WO 0196581-A/1					
	PD 20-DEC-2001					
	PF 15-JUN-2001 WO 2001JP005096					
	PR 15-JUN-2000 JP 00P 190466					
	PI ATSUHIKO SHIMMO,KO KATO,YASUHIRO YAMADA,TAKUYA NIHIRA,TAKUYA SHINDO					
	PC C12N15/82,A01H5/00,C12N5/10					
	CC Method for induction of gene expression in plants and plants thereof					
	CC Method for induction of gene expression in plants and plants thereof					
FT	Key	Location/Qualifiers				
	CDS	(1)..(699).				
FEATURES	Location/Qualifiers					
source	1..699					
	/organism="Streptomyces virginiae"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:1961"					
ORIGIN						
	Query Match	100.0%;	Score	699;	DB	6;
	Best Local Similarity	100.0%;	Pred. No.	5.9e-75;		
	Matches	699;	Conservative	0;	Mismatches	0;
					Indels	0;
					Gaps	0;


```
Db 874 CAGGAACGGGGGAGGTCTTCGCGACGTAACCCGAAAGAACCGCGGACTTCATCGTC 933
Qy 481 GGCTGCTTACCGGGGTCAGGGCGGTCTCCCGGGTCACTCCGACCGCCAGGACCTCGGC 540
Db 934 GGCTGCTTACCGGGGTCAGGGCGGTCTCCCGGGTCACTCCGACCGCCAGGACCTCGGC 993
Qy 541 CACCGATCTCGTGATGTGAACACCGTCTGCCAGATCTGCGGGGTCATGCTG 600
Db 994 CACCGATCTCGTGATGTGAACACCGTCTGCCAGATCTGCGGGGTCATGCTG 1053
Qy 601 ACCTGATCGAACCAGGAGCGATCGGAAGTCCGCGCGCGCGGCGCGCC 660
Db 1054 ACCTGATCGAACCAGGAGCGATCGGAAGTCCGCGCGCGCGGCGCGCC 1113
Qy 661 GAGGCTCGGAGGCTCCGAGGCGGCTCCGACGAGTAG 699
Db 1114 GAGGCTCGGAGGCTCCGAGGCGGCTCCGACGAGTAG 1152

RESULT 3
AB001683 2234 bp DNA linear BCT 22-AUG-1997
LOCUS Streptomyces sp. gene for FarX, FarA, complete cds.
DEFINITION Streptomyces sp.
ACCESSION AB001683
VERSION AB001683.1 GI:2342428
KEYWORDS FarX, FarA.
SOURCE Streptomyces sp.
ORGANISM Streptomyces sp.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE
AUTHORS Waki,M., Nihira,T. and Yamada,Y.
TITLE Cloning and characterization of the gene (farA) encoding the
receptor for an extracellular regulatory factor (IM-2) from
Streptomyces sp. strain FRI-5
JOURNAL J. Bacteriol. 179 (16), 5131-5137 (1997)
MEDLINE 97405912
PUBMED 9260956
REFERENCE 2 (bases 1 to 2234)
AUTHORS Waki,M.
TITLE Direct Submission
JOURNAL Submitted (07-MAR-1997) Miyoko Waki, Osaka University, Graduate
School of Engineering, Department of Biotechnology, Yamadaoka 2-1,
Suita, Osaka 565, Japan (E-mail:waki.yam@stu.bio.eng.osaka-u.ac.jp,
Tel.:+81-6-879-7433, Fax:+81-6-879-7432)
FEATURES
source
1..2234
/organism="Streptomyces sp."
/mol_type="genomic DNA"
/strain="FRI-5"
/db_xref="taxon:1931"
440..1315
/gene="farX"
440..1315
/gene="farX"
/codon_start=1
/transl_except=(pos:440..442,aa:Met)
/transl_table=11
/product="FarX"
/protein_id="BAA21858.1"
/db_xref="GI:2342428"
/translations="MHRTSTAQVLLTMDQDDAFSTARWPLSHAFPTVPDGYG
DPMCAETRIQAVLLGHAEFVFGHOFVLDLSVVRPELLRVLGPATVDLAI
CWEIKRAGELSLGYEAVVRDQGVATGCRASVTCTSPAVQRIPEHVLTPHRL
PLTAPAPQSVARLSPTDVLSPDLRENWQLRVDNHPVLFHWDHPGVMVMAA
RQAASALGRSPFMPPLGVAGEFKRYVELDAPCVIESERLFDQVFGAEVVRVTGQNG
ELTFVGTVTASSYGY"
1500..2165
/gene="farA"
1500..2165
/gene="farA"
/codon_start=1
```

```
/transl_except=(pos:1500..1502,aa:Met)
/transl_table=11
/product="FarA"
/protein_id="BAA21859.1"
/db_xref="GI:2342430"
/translations="NAEOVRAIRTRQAILSAARVFDERGYOAAITSEILTAVGVTKG
ALYFHFQSKEDLAQVILTAQNEIDLILPERPAKQEVVDVAVMLHTRLRPMVRAGVR
LSLDVAGGLDRSAPFRNWDKFTDLLEKAQAGELLPHVPAETRADVITGAYGVQS
MSQALTEHQDLQGRVNLRLHLMPSIAQPSVLASLHLSRAEVEVYLEARQAREAD
EBD"

ORIGIN
Query Match 38.2%; Score 267.2; DB 1; Length 2234;
Best Local Similarity 63.0%; Pred. No. 2.6e-23;
Matches 413; Conservative 0; Mismatches 243; Indels 0; Gaps 0;

Qy 33 ACAGGAACGGGGCGTCCGACCGCGGCGGATCGTGGGGGAGCGGCTCGGTCTTCGA 92
Db 1508 ACAGGTCCGAGCATCCGACCGGCGGCGGATCTTGAGCGGCGCGGAGGTGTTCA 1567
Qy 93 CGAGTACGGGTTGAGGGCGCCACAGTGGCAGAGATCTCTCGGGGCTTCGTCACCAA 152
Db 1568 CGAACGGGGCTACCGAGCGGCGCACCATCTCGGAGATCTCTCACGTCGCGGGGTGACCAA 1627
Qy 153 GGGCGGATGTACTTCCACTTCGCTTCCAAGGAGAGCTGGCCCGGGCGGTGCTGCCGA 212
Db 1628 GGGCGGCTGTACTTCCACTTCAGTCCAAGGAGACCTGGCCCGGCGGTCTCTACCGC 1687
Qy 213 GCAGACCTGCGACGTGGCGGTGCGGAATCCGCTCCAAAGCGCAGAGAACTGGTAGACCT 272
Db 1688 GCAGACGAGGATCTTCTCTCTGAGGCGCCCGCCAACTCCAGAGGTGTTGATGC 1747
Qy 273 CACCATGTGTTGCGCCACCGGATGTGCACGATCGATCTCTCGGGCGGCGACGGGCT 332
Db 1748 CGTCATGTGTCACCCACCGCGCTCGGACCAACCCCATGTTCCGCGCGGCGTCCGGCT 1807
Qy 333 CGCACTGAGACCGGGCGGTGGACTTCTCGACGCCAACCCGTTGCGGAGTGGGGCGA 392
Db 1808 GAGCTTCGACGTGAACCGCGGGGGGTTCGACCGCGGCTTCGCAACTGGGTGCGA 1867
Qy 393 CATCTCGGCCAGCTCTCTCGGAGGACACAGGAACGGGGGAGGTCTTCCGACACGTAA 452
Db 1868 CAGTTACCGACCTCTCGGAGAGCGCCAGGCCAGGCCAGGAACTGTCGCCACGTGT 1927
Qy 453 CCGCAAAAGACCGGCGATTCTATCTCGGTGCTTCAACGGGCTTCAGGCGGTCTCCCG 512
Db 1928 ACCGGCGGAGACCGCGCGGTCTATCACCGGTCTCTACGGCGGCTCCAGTCCATGTCCA 1987
Qy 513 GGTCACTCGGACCGCGGACCTCGGCCACCGATCTCGGTGATCTGGAACACCGTCT 572
Db 1988 GCGCTCACCGAGCACCGGACCTCGGGCAGCGGGTCAACGCGCTGTCGCCACCTCAT 2047
Qy 573 GCCCAGCATCTGCGCGGCTCCATGCTGATCGATCGAATCGAATCGGAGAGCGGATCGG 632
Db 2048 GCGGACATCGCCAGCCCTCGGTCTCTCGCTCCCTGACCTCGGCGAGAGCGGGCGGA 2107
Qy 633 GAAGTTCGCGCGGCGCGGAGCGCCCGAGGCTGCGGAGGCTTCGAGCGCCGCT 688
Db 2108 AGAGGTCTACCTCGAAGCCCGCGAGCTGCGCCCGGAGCAGCGGAGGAGAACT 2163

RESULT 4
SC0007731 6727 bp DNA linear BCT 24-AUG-1998
LOCUS Streptomyces coelicolor scbr gene, scbA gene, ORFs A,B,X & Z.
DEFINITION Streptomyces coelicolor scbr gene, scbA gene; scbr gene.
ACCESSION AJ007731
VERSION AJ007731.1 GI:3425857
KEYWORDS gamma-butyrolactone binding protein; scbA gene; scbr gene.
SOURCE Streptomyces coelicolor A3(2)
ORGANISM Streptomyces coelicolor A3(2)
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE
1
```


Db 3341 TCATGACACAGAGCGGCTCTGATCGCCGAGGACCCCTTCGTCGTGGACGAG 3400
Qy 394 ATCTGCGCCAGCTCTCGCGGAGGACACGAGGACGCGGGGAGGTGCTTCGCGACGTGAAC 453
Db 3401 ACACCTCTGAAGCTGCTGAACACGAGGACGAGGACGAGTGTGCTGCCCCATGTGTC 3460
Qy 454 CGGAAAAGACCGGACCTCATCTGCTGGCTGCTTACCGGGTCCAGCGGTCTCCCGG 513
Db 3461 ACCACGACTCGGCCGATCTTACGTTGGGCACGCTTCGCGGGATACAGGTCTGTGTC 3520
Qy 514 GTCACTTCGACCGCGCAGACCTCGCCACCGGATCTCGTGATGGAACACGCGTGTG 573
Db 3521 ACAGTACGAGCTACACGACCTCGAACCGCTGACCGCTGCTGCGAGACACATCTCTG 3580
Qy 574 CCAGACGCTGCGCGCTCCATCTGCTGACCTGATCGAAACCGCGGAGGACGATCGGG 633
Db 3581 CCGGCATCGCGTTCCTCCGCTGCTGCGCGCTGCTGCTCCGAGGCGCGGACGA 3640
Qy 634 AGGTCGCGCGG 646
Db 3641 CGCCTCGCGCGC 3653

RESULT 5
SC0939127/c
LOCUS SC0939127 290850 bp DNA linear BCT 11-FEB-2003
DEFINITION Streptomyces coelicolor A3(2) complete genome; segment 24/29.
ACCESSION AL93127 AL023496 AL031861 AL031155 AL031182 AL031232 AL035161
AL035205 AL035206 AL132824 AL512902 AL590982 AL591083 AL591084
AL645882
VERSION AL939127.1 GI:24429552
KEYWORDS Streptomyces coelicolor A3(2)
SOURCE Streptomyces coelicolor A3(2)
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE 1
AUTHORS Bentley,S.D., Chater,K.F., Cerdeno-Tarraga,A.M., Challis,G.L., Thomson,N.R., James,K.D., Harris,D.E., Quail,M.A., Kieser,H., Harper,D., Bateman,A., Brown,S., Chandra,G., Chen,C.W., Collins,M., Cronin,A., Fraser,A., Goble,A., Hidalgo,J., Hornsby,T., Howarth,S., Huang,C.H., Kieser,T., Larke,L., Murphy,L., Oliver,K., O'Neill,S., Rabinowitch,E., Rajandream,M.A., Rutherford,K., Rutter,S., Seeger,K., Saunders,D., Sharp,S., Squares,R., Squares,S., Taylor,K., Warren,T., Wietzorrek,A., Woodward,J., Barrell,B.G., Parkhill,J. and Hopwood,D.A.
TITLE Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2)
JOURNAL Nature 417 (6885), 141-147 (2002)
MEDLINE 21996410
PUBMED 12000953
REFERENCE 2 (bases 1 to 290850)
AUTHORS Bentley,S.D.
JOURNAL Direct Submission
TITLE Submitted (09-MAY-2002) Submitted on behalf of the Streptomyces sequencing team, Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: sdb@sanger.ac.uk
COMMENT On or before Oct 30, 2002 this sequence version replaced gi:20520661, gi:20520749, gi:20520754, gi:20520755, gi:20520688, gi:20520689, gi:20520815, gi:20520683, gi:20520917, gi:20520866, gi:20520864.

FEATURES
source Location/Qualifiers
1. .290850
/organism="Streptomyces coelicolor A3(2)"
/mol_type="genomic DNA"
/strain="A3(2)"
/db_xref="taxon:100226"
97. .1101
/gene="SCO6257"
/notes="synonym: SCAH10.22"
97. .1101
/gene="SCO6257"
CDS
/gene="SCO6257"

/note="SCAH10.22, probable ABC transport system sugar binding lipoprotein, len: 334 aa; similar to TR:CA41563 (EMBL:AL049727) Streptomyces coelicolor putative secreted solute binding protein, 337 aa; fasta scores: opt: 1003 z-score: 1130.9 E(0); 49.8% identity in 325 aa overlap, to SW:R8SB_BACSU (EMBL:Z92953) Bacillus subtilis D-ribose-binding protein precursor RbsB, 305 aa; fasta scores: opt: 206 z-score: 238.7 E(0); 6.3e-06; 23.0% identity in 235 aa overlap and to TR:CA41563 (EMBL:SC981) Streptomyces coelicolor SC981.17c, 337 aa; fasta scores: opt: 1003 z-score: 1045.2 E(0); 49.8% identity in 325 aa overlap. Contains match to Pfam entry PF00532 Peripla_BP_like, periplasmic binding proteins and LacI family and a match situated in the correct position to Prosite entry PS00013 Prokaryotic membrane lipoprotein lipid attachment site. Contains also possible N-terminal region signal peptide sequence"
/codon_start=1
/transl_table=11
/product="putative ABC transport system sugar binding lipoprotein"
/protein_id="CAB60176.1"
/db_xref="GI:6273662"
/translation="MARPTWVGIALGALSUSLAGCSSTGCKRAEDARKAASAGRA AVDPRTWPMITHSGDGDFTWDIVQSGAEQAQVNDINFLYSHDREAQQQQLVNAADKVDGIIIVTLAKPDAMKSLARAHKAGIPVITNSGSEESKEFGALTHVGODETIA GEAVEELNEHQEQAVCVLHEQNVGHEQSCDGVKTFDQKQRLYVNGTSPDVQS AIEALQTDKSDVAVVTLGAPVADYAVAKQAGKAEIDTDLNAAKVAAGLADGTLG FAVDQPPYLGVEYAVDLLWLYKNADVLGGGRPVLTGPQIITKDDAAALADYTKRGT R"
misc_feature 133. .165
/gene="SCO6257"
/note="PS00013 Prokaryotic membrane lipoprotein lipid attachment site"
241. .966
/gene="SCO6257"
/note="Pfam match to entry PF00532 Peripla_BP_like, periplasmic binding proteins and LacI family, score 13.80, E-value 2.1e-06"
1098. .2138
/gene="SCO6258"
/note="synonym: SCAH10.23"
1098. .2138
/gene="SCO6258"
/note="SCAH10.23, probable ABC transport system sugar permealase, len: 346 aa; similar to TR:O68120 (EMBL:AF010496) Rhodobacter capsulatus ribose transport system permealase protein RbsC, fasta scores: opt: 445 z-score: 508.0 E(0); 6.4e-21; 30.5% identity in 351 aa overlap and to SW:R8SC_ECOLI (EMBL:L10328) Escherichia coli ribose transport system permealase protein RbsC, len: 321 aa; fasta scores: opt: 286 z-score: 329.9 E(0); 5.3e-11; 30.1% identity in 329 aa overlap. Contains possible hydrophobic membrane spanning regions"
/codon_start=1
/transl_table=11
/product="putative ABC transport system sugar permealase"
/protein_id="CAB60177.1"
/db_xref="GI:6273663"
/db_xref="GOA:Q9RKT4"
/db_xref="UniProt/TREMBL:Q9RKT4"
/translation="MNAQVKAADRIILQTSRLRLKLRPELGSVVGAATAVFVFFAFF ADSFLHAASLSTLVAASTGIMAVFALLMGFFDLSAGVMVTSALSMFVSQMTANVWVGVVSLVTLAIGAFNGFMLTKLPSTITLTFLMLGLNGLFKLVNVTGTSKSIADMEGPPSQDVFASITIGGVGFKVITLWMLAVAVASWILLRAGNWIIFAVGKNDAARAVGVPAVKTKGLYMGVGFAGWISQHLIFSVDVVGSGVGVELIYI TAIVIGGCLITGTVGAVSGSAVGAFIFGTSKGIIVFAEWNPDWFRFFLGMALLATLL NAWVRKRAEATK"
RBS 2124. .2127
/gene="SCO6258"
2135. .2926
/gene="SCO6259"

CDS	/note="synonym: SCAH10.24" 2135..2926 /gene="SCO6259" /note="SCAH10.24, probable ABC sugar transport ATP binding protein, len: 263 aa; similar to TR:O68121 (EMBL:AF010496) Rhodobacter capsulatus ribose transport ATP-binding protein, 305 aa; fasta scores: opt: 612 z-score: 706.1 E(): 5.9e-32; 45.6% identity in 217 aa overlap, to SW:RBSA_ECOLI (EMBL:M13169) Escherichia coli, ribose transport ATP-binding protein RbsA, 501 aa; fasta scores: opt: 509 z-score: 585.6 E(): 3e-25; 31.5% identity in 254 aa overlap and to TR:CAH15787 (EMBL:SC7B7) Streptomyces coelicolor St7B7.07, 260 aa; fasta scores: opt: 592 z-score: 521.2 E(): 1e-23; 43.1% identity in 246 aa overlap. Contains Pfam match to entry PF00005 ABC tran, ABC transporter and Prosite PS00017 ATP/GTP-binding site motif A (P-loop) and PS00211 ABC transporters family signature" /codon_start=1 /transl_table=1 /product="probable ABC sugar transport ATP binding protein" /protein_id="CAB60178.1" /db_xref="GI:6273664" /db_xref="GOA:Q9RKT3" /db_xref="UniProt/TREMBL:Q9RKT3" /translation="WTRNEDRTALVELSGVSKNYGNVRALEGVSLVHAGEITCVLGD NGAGSTLKIITAGLHQHDGTLSDIGETRLSSPREALDRGIATVYQDLAVVPLMPV WRNPLGSPRKGVAPEKRMVDHMRTHASLRNMGIDLRVDQPIGTLSGERQCV ATARAVYFGAKVLVDLEPTAALGVKSGVVLKVAARDQGLGVLIITHNPHAYLVG DRFVLKRGTVANQTRDVBTLDELTFQWAGTELDLRLHELRG" 2240..2815 /gene="SCO6259" /note="Pfam match to entry PF00005 ABC tran, ABC transporter, score 168.70, E-value 9.6e-47" 2261..2284 /gene="SCO6259" /note="PS00017 ATP/GTP-binding site motif A (P-loop)" 2588..2632 /gene="SCO6259" /note="SCO0211 ABC transporters family signature" 3042..4190 /gene="SCO6260" /note="synonym: SCAH10.25" 3042..4190 /gene="SCO6260" /note="SCAH10.25, possible sugar kinase, len: 382 aa; similar to SW:GLK_STRCO (EMBL:X65932) Streptomyces coelicolor glucokinase (EC 2.7.1.2) (glucose kinase) Glk, 317 aa; fasta scores: opt: 1022 z-score: 1142.9 E(): 0; 46.6% identity in 311 aa overlap and to TR:CAB51974 EMBL:SC6E10 Streptomyces coelicolor SC6E10.20c, 317 aa; fasta scores: opt: 1022 z-score: 1035.4 E(): 0; 46.6% identity in 311 aa overlap. Contains a match to Pfam entry PF00480 ROK, ROK family and Prosite entry PS01125 ROK family signature" /codon_start=1 /transl_table=1 /product="putative sugar kinase" /protein_id="CAB60179.1" /db_xref="GI:6273665" /db_xref="GOA:Q9RKT2" /db_xref="UniProt/TREMBL:Q9RKT2" /translation="MSTVRDFTAPIGSRAPVLRTVGTRERRSHLTAPRPVTVGIDIG GTKWAGVVDADGNILEKLTETPDKSPKVEDTIVELVLDLSDRDHVAVGIAA GVVDAADRNLVAPLHSLWRNEPLRDIAGLAPVPLVDNDANTAAWRFAGAGED HLVTITLGTIGGAILDQGVKRGYGVAGFGHMVQVPGHRCPCNGRNGWEQVSSG NALVREARELAADSPVAYGIIIEHVKSGIDITGPMITELAREGAMCVELLQIDQGW LGVGLANLAALDPSCFVIGGVSAADDLLIGFAPDAFKRLTGRTGYRPEARIVRAQL GPAGMVGAAADLSRLVARRFRKRRRVERERYERYAEARESESL" 3159..3704 /gene="SCO6260" /note="Pfam match to entry PF00480 ROK, ROK family, score		misc_feature	215.50, E-value 5.4e-63" 3540..3623 /gene="SCO6260" /note="PS01125 ROK family signature" 4244..4247 4259..4768 /gene="SCO6261"	misc_feature	Query Match 32.8%; Score 229; DB 1; Length 290850; Best Local Similarity 60.8%; Pred. No. 2.5e-19; Matches 373; Conservative 0; Mismatches 240; Indels 0; Gaps 0;	QY	34	CAGGAACGGGCGTCCCGACGCGGAGCGATCGTGGGCGACGCGCTCGTCTTCGAC	93
					Db	11966	CAGGACCGGGCGATCCGACGCGGAGACGATCTCTGGAGCGCGCGAGGCTTCGAG	11907		
					QY	94	GAGTACGGGTTGAGCGCGCCACAGTGGCAGAGATCTCTCGGGGGCGCTCGGTCAACAAG	153		
					Db	11906	AAGCAGGGCTACCAAGCTGCCACGATCAGGAGATCTCAAGGTGGCGGGTGACCAAG	11847		
					QY	154	GGCGCGATGTACTTCCACTTCGCTTCCAAAGGAAGAGCTGGCCCGCGGTGCTGGCCGAG	213		
					Db	11846	GGAGCGCTTACTTCCACTTCCAGTCCAGGAAAGAACTGGCGTGGCGGCTTTCGACGCC	11787		
					QY	214	CAGACCTGACGTCGCGGTGCCGAATCCGCTCCAAAGGCGCAGGAACTGGTAGACCTC	273		
					Db	11786	CAGGAACACACACAGCGCGTTCCGAGCAACCCCTCCGCTGCAAGAACTCATCGACATG	11727		
					QY	274	ACCATGCTGTCGCCACCGGATGCTGCAGATCCGATCTCTGGGCGGCGGACGCGGCTC	333		
					Db	11726	GGCATGTGTCTGTGTCACCGCTTGGCGACGAACTGTCGGCGCGCGGTGCGGCTC	11667		
					QY	334	GCATCTGACACAGGGCGGTGGACTTCTCCAGCGCAACCCGTTCCGGCGAGTGGGGCGAC	393		
					Db	11666	TCCATGGACAGGCGGACCGTCTCGATGCCGAGGACCTTCGCTCGCTGGCAGCAG	11607		
					QY	394	ATCTCGGCGGCTCTCTGGGAGGACAGCAAGAAACGGGGGAGTGTCTTCGCACTGAAC	453		
					Db	11606	ACACTCTGAAAGTCTGTAACAGGCCAAGAGAACGGTGTCTGCTGCCCATGTGCTC	11547		
					QY	454	CGMAAAGACCGGCGACTTCATCGTGGCTGCTTACCGGGCTCCAGCGGTCTCCCGG	513		
					Db	11546	ACCACCGACTCGGCGGATCTCTACGTGGGACGCTTCGCGGGGATACAGTGTGTGCCAG	11487		
					QY	514	GTCACCTCCGACCGCAGGACCTCGGCCACCGGATCTCGGTGTATGTGGAACACCGTGTG	573		
					Db	11486	ACGGTCAGCGACTTACACGAGCTTCGAAACACCGCTACGCGCTGCTGCAGAAACATCTG	11427		
					QY	574	CCGACATCGTCCCGCGTCCATGCTGACCTGGATCGAAACCGGCGGAGAGCGGATCGGG	633		
					Db	11426	CCCGCATCGCGTTCCTCCGCTGCTGCTGCGCGCGCTCGATCTCTCCGAGGAGCGCGGAGCA	11367		
					QY	634	AAGGTCCGCGCGG 646			
					Db	11366	CGCTCGCGCGG 11354			
					RESULT 6					
					STWAFR					
					LOCUS		1328 bp	DNA	linear	BCT 10-FEB-1999
					DEFINITION	Streptomyces griseus DNA for A-factor receptor, complete cds.				
					ACCESSION	D49782				
					VERSION	D49782.1 GI:1313911				
					KEYWORDS	A-factor receptor.				
					SOURCE	Streptomyces griseus				
					ORGANISM	Streptomyces griseus				
					REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				
					AUTHORS	1 (bases 1 to 1328)				
					TITLE	Streptomycetaceae; Streptomyces.				
					JOURNAL	Streptomyces griseus J. Bacteriol. 177 (21), 6083-6092 (1995)				

MEDLINE 96042082
PUBMED 7592371
REFERENCE 2 (bases 1 to 1328)
AUTHORS Onaka,H.
TITLE Direct Submission
JOURNAL Submitted (18-MAR-1995) Hiroyasu Onaka, The University of Tokyo,
Department of Biotechnology; Yayoi, Bunkyo-ku, Tokyo 113, Japan
(E-mail:aa57093@hongo.ecc.u-tokyo.ac.jp, Tel:03-3812-2111(ex:5147),
Fax:03-5802-2931)
COMMENT On May 11, 1996 this sequence version replaced gi:1089808.
FEATURES
source
1..1328
/organism="Streptomyces griseus"
/mol_type="genomic DNA"
/strain="IFO13350"
/db_xref="taxon:1911"
299..1129
/codon_start=1
/transl_table=11
/product="A-factor receptor"
/protein_id="BAA08617.1"
/db_xref="GI:1742915"
"translations="MAKQARAVOTWRSIVDAASVFDYGERAAISEILRAKVTG
ALYHPASKEAIAQATMDQTSVEFEQSGPLOSIVDGSQOPAFALHNSVARAKTR
LSEGVPLGCPHPGWIDATARMLELQERGEVFIQIDPMVSAKLIIVASFQIQIUV
EADSGRADUREQVAEMRHLPISIAHPGVIAHIKPEGRVDLAAQAREKAREBEQEAR
AAEAKAGSDPTSEGGTRSGSLRGSGRGRPRAGVTGDEDEEPAGVAAAGGIVA
"

Query Match 30.6%; Score 213.6; DB 1; Length 1328;
Best Local Similarity 58.7%; Pred. No. 8.1e-17;
Matches 391; Conservative 0; Mismatches 269; Indels 6; Gaps 1;
QY 34 CAGGAACGGGCGCTCCGACGCGGCGGCGATCGTGGGCGGCGCTCGGTCTTCGAC 93
DB 308 CAGGCTCGCGCATGCAGACGTGGCGTGCATCGTGTGATGCCGCGGAGTGTTTCGAC 367
QY 94 GAGTACGGGTTCAGAGCGGCCACAGTGGCAGAGATCCTCTCGCGGCGCTCGGTCAACAAG 153
DB 368 GACTACGGCTACGAGCGTGGCGGCATCTCGGAGATTCTGGCGCGCGCAAGGTCAACAAG 427
QY 154 GGGCGGATGTACTTCACTTCGCTTCCAGGAGAGAGCTGGCGCGGCGGCTGGTGGCGGAG 213
DB 428 GGGGCGCTTGACTTCACTTCGCTTCCAGGAGGCGCATCGCCCGGCGGACTCATGAGCGAG 487
QY 214 CAGACCGCTGCACGTGGCGGTCGCGAATCCGGTCCAAAGCGCAGGAACCTGGTAGACCTC 273
DB 488 CAGAGCTCAGGTGGAGTTCGAGCAGGAGGCTTCGCGCTTCAGTCCCTGGTGGAGCGG 547
QY 274 ACCATGCTGTGCCCAACGCGCATCTGTGACGATCCGATCCGATCTCGCGGCGGCGCACGCGGCTC 333
DB 548 GGCAGCAGTTCGCTTTCGCGCTTCGCGCAAACTCGATGGCGCGCGCGGTACACGCGCTC 607
QY 334 GCATGGACAGGGGGGGGTGGACTTCTCCGACGCCAACCCGTTCCGCGAGTGGGCGGAC 393
DB 608 TCCATCGA-----GGCGGTCTTCCTCGCGGCGCGCGCACCCCTGGGCGGAGTGGATCGAC 661
QY 394 ATCTGCGCCAGCTCTCTGGCGGAGGACAGGAACGGGGGAGTGCTTCGCGCAGCTGAAC 453
DB 662 GCACGCGCCCGGATGCTGGAGTGTGGCCAGAGCGCGGCGAGGTGTTCGCCAGATGCGAC 721
QY 454 CCGAAAAAGACCGGCGCATTTTCATCGTGGGTGTCTTCAACCGGGGTCCAGGCGGCTCTCCCGG 513
DB 722 CCGATGGTGTTCAGCAAAATCATCGTCTTCGTTTCAACCGGTATCCAGCTGGTCTCGGAG 781
QY 514 GTCACTCCGACCGCGAGGACCTTCGGCCACCGGATCTCGTGTATGTGGAAACCAACGCTGCTG 573
DB 782 GCCGATTCGCGCGCGGCGGATCTTCGCGAGCAGGTGGCGGAGATGTGGCGGACATCTCTG 841
QY 574 CCCAGCATCTGCGCGGCTCCATCTGCTGACCTGATCGAAACCGCGGAGGACGCGGCTCGG 633
DB 842 CCGTCGATCCGCCACCCCGGGGTTCATCGCCCAATCAACCGGAGGCGCGGTGGATCTG 901

ORIGIN
QY 634 AAGGTCG 693
DB 902 GCGGCGGAG 961
QY 694 GAGTAG 699
DB 962 GCAAG 967

RESULT 7
AY338477/c 33064 bp DNA linear HTG 07-OCT-2004
LOCUS Streptomyces ambofaciens clone cosmid F6 strain ATCC 23877, ***
DEFINITION SEQUENCING IN PROGRESS ***, 2 ordered pieces.
ACCESSION AY338477 AY338478
VERSION AY338477.2 GI:53851167
KEYWORDS HTG; HTGS_PHASE2
SOURCE Streptomyces ambofaciens
ORGANISM Streptomyces ambofaciens
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomycetes.
REFERENCE 1 (bases 1 to 33064)
AUTHORS Pang,X.,Aigle,B., Girardet,J.M., Mangenot,S., Pernodet,J.L.,
Decaris,B. and Leblond,P.
TITLE Functional Angucycline-Like Antibiotic Gene Cluster in the Terminal
Inverted Repeats of the Streptomycetes ambofaciens Linear Chromosome
JOURNAL Antimicrob. Agents Chemother. 48 (2), 575-588 (2004)
PUBMED 14742212
REFERENCE 2 (bases 1 to 33064)
AUTHORS Aigle,B.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-2003) Genetique et Microbiologie, UMR INRA 1128,
IFR 110, Universite Henri Poincare, Faculte des Sciences et
Techniques, Boulevard des Aiguillettes, BP 239,
Vandoeuvre-les-Nancy 54 506, France
REFERENCE 3 (bases 1 to 33064)
AUTHORS Aigle,B.
TITLE Direct Submission
JOURNAL Submitted (06-OCT-2004) Genetique et Microbiologie, UMR INRA 1128,
IFR 110, Universite Henri Poincare, Faculte des Sciences et
Techniques, Boulevard des Aiguillettes, BP 239,
Vandoeuvre-les-Nancy 54 506, France
REMARK
COMMENT Sequence update by submitter
On Oct 7, 2004 this sequence version replaced gi:39748104.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 13715: contig of 13715 bp in length
* 13716 13815: gap of unknown length
* 13816 33064: contig of 19249 bp in length.
FEATURES
source
1..33064
/organism="Streptomyces ambofaciens"
/mol_type="genomic DNA"
/strain="ATCC 23877"
/db_xref="ATCC:23877"
/db_xref="taxon:1889"
/clone="cosmid F6"
239..1324
/gene="alpH"
239..1324
/gene="alpH"
/note="similar to pfam PF00891 O-methyltransferase"
/codon_start=1
/evidence=not experimental
/transl_table=11

Db 32858 CGGACGACGAGCGATCCGTACCCAGTTGTCATCTCGATCGAGCGTGGCGGCGAGGTGTTCC 32799

Qy 91 GACGAGTACGGTTCAGAGCGCCGACAGTGGCAGAGATCTCTCGCGGGCTCGGTCAACC 150
|||||
Db 32798 GACTCCCATGGCTACGAGGCGCGACCACTCGGGAGATCTCTAGACGGGCGGTGTCAACC 32739
|||||

Qy 151 AAGGGCGCATGTACTTCCATTTGGTTCCTCAAGGAGAGTGGCCGCGCGTCTGGGCC 210
|||||
Db 32738 AAGGGCGCGTGTATTTCATTTCCCTTCCCAAGCAGCTCTGGCGAGGGGGTGTGGAG 32679
|||||

Qy 211 GACGACACCTGTCACGTCGGCGGTGGCGGAATCCGGCTCCAAAGCGCAGGAACCTGGTAGAC 270
|||||
Db 32678 CACGAGTTCTCCGTGATCCGGTGGCGCGCGTCCGTGAGGCTCAGGAGTTCTGGGAC 32619
|||||

Qy 271 CTCACCATGTCGTGCGCCACAGCATGCTGCACGATCCGATCTCTGGCGGCGGACCGCG 330
|||||

Db 32618 ACGGGTCTGATCTGGCGCTACCGCATCGCGCGTGTATCCGCTGGTGGCGCGGTCCGACG 32559
|||||

Qy 331 CTCGCACTGACACGAGGGCGGTGGACTTCTCGACGCCAACCCGTTCCGCGAGTGGGC 390
|||||

Db 32558 CTGAGTCTGGAGCAGGAACCTGCTGGGAGTACGGCTGTCGGCGATCCGGCAGTGGATC 32499
|||||

Qy 391 GACATCTGCGCCAGCTCTTGGCGGAGGACACGAAACCGGGGAGGTGCTTCCGACGCTG 450
|||||

Db 32498 GGGCGCTCCGAGTGTCTGGTGGCGCGGAGGAGGAGTGTCTTCCGCGATGG 32439
|||||

Qy 451 AACCGAAAGACCGCGCATCTTCATGTCGCTGCTTCCACCGGCTCCAGGCGGTCTCC 510
|||||

Db 32438 GTCCGCGCCGAGAGTGGTGGTGTCTCGCGCGGTGGACCGGTACGCGAGCTCTACTCG 32379
|||||

Qy 511 CGGCTCACCTCCACGCGCAGGACCTCGGCCACCGATCTCGTGTATGTGGACACGCTG 570
|||||

Db 32378 CAGATCTGTCGGCGCGGAGGACCTGGAGGAGCGGTGTGCTCTTCCGCGACCTG 32319
|||||

Qy 571 CTCGCCAGCATCTGCGCGCTCATGCTGCTGACCTGATCGAAACCGCGAGGACCGATC 630
|||||

Db 32318 CTGCCAGCATCCCGTGGCGCGGTGCTCAGCAGTCTGGAGATCACCCAGCGCGGCG 32259
|||||

Qy 631 GGAAGGTTCGCGCGCGCGCGGAGCGCGCGAGGCTCGGAGCGCTCCGAGGCGCGCTCC 690
|||||

Db 32258 GCCCGCTCGTTTCGCGAGCGAGCGGTGCTCGAGGCGCGCGCGGAGCGGAGCTCC 32199
|||||

Qy 691 GACGAGTAG 699
|||||

Db 32198 CGCAGGAG 32190

RESULT 8
AB088224/c

LOCUS Streptomyces rochei plasmid pSLA2-L DNA linear BCT 11-JUN-2003
DEFINITION Streptomyces rochei plasmid pSLA2-L DNA, complete sequence.
ACCESSION AB088224
VERSION AB088224.1 GI:30698345

KEYWORDS

SOURCE Streptomyces rochei
ORGANISM Streptomyces rochei
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Streptomyces.

REFERENCE 1
Kinashi, H., Sugino, H., Sasaoka, A., Mori, E., Fujii, S., Shinkawa, H.,
Nimi, O. and Kinashi, H.
Physical mapping of the linear plasmid pSLA2-L and localization of
the eryA and actI homologs
Biosci. Biotech. Biochem. 62, 1992-1997 (1998)

TITLE Identification of two polyketide synthase gene clusters on the
linear plasmid pSLA2-L in Streptomyces rochei
Gene 246 (1-2), 123-131 (2000)

JOURNAL 20231737
MEDLINE 10767533
PUBMED

AUTHORS Hiratsugu, K., Mochizuki, S. and Kinashi, H.

TITLE Cloning and analysis of the replication origin and the telomeres of
the large linear plasmid pSLA2-L in Streptomyces rochei
Mol. Gen. Genet. 263 (6), 1015-1021 (2000)

JOURNAL 20408175
MEDLINE 10954087
PUBMED

REFERENCE 4
Mochizuki, S., Hiratsugu, K., Suwa, M., Ishii, T., Sugino, F., Yamada, K.
and Kinashi, H.
The large linear plasmid pSLA2-L of Streptomyces rochei has an
unusually condensed gene organization for secondary metabolism
Mol. Microbiol. 48 (6), 1501-1510 (2003)

TITLE Cloning and analysis of the replication origin and the telomeres of
the large linear plasmid pSLA2-L of Streptomyces rochei
Mol. Microbiol. 48 (6), 1501-1510 (2003)

JOURNAL 12791134
MEDLINE 22676866
PUBMED

AUTHORS Mochizuki, S., Hiratsugu, K. and Kinashi, H.
Direct Submission
Submitted (15-JUL-2002) Haruyasu Kinashi, Hiroshima University,
Department of Molecular Biotechnology, Graduate School of Advanced
Sciences of Matter; 1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima
739-8530, Japan (E-mail: kinashi@hiroshima-u.ac.jp,
Tel: 81-824-24-7869, Fax: 81-824-24-7869)

TITLE The nucleotide sequence has been determined by using restriction
fragments and nested deletion fragments of the ordered cosmid
library of pSLA2-L.
pSLA2-L is 210,614 bp long with G+C content of 72.8 % and carries
143 ORFs.

COMMENT Gene prediction was based on the unique codon usage in Streptomyces
(Bibb et al., Gene 30:157-66 (1984)) using the FramePlot program of
Ishikawa and Hotta (FEMS Microbiol Lett 174: 251-253 (1999)) as
implemented at
http://www.nih.go.jp/~jun/cgi-bin/frameplot.pl. Where possible we
chose an initiation codon (atg, gtg, ctg or ttg) which is preceded
by an upstream ribosome binding site sequence (optimally 5-13 bp
before the initiation codon). If this could not be identified we
chose the most upstream initiation codon.
Three-fourths of the pSLA2-L DNA are occupied by biosynthetic genes
(three PKS gene clusters for lankacidin, lankamycin and an unknown
type II polyketide, a carotenoid biosynthetic gene cluster, many
regulatory genes and others). The range of each biosynthetic gene
cluster has been deduced by comparison with similar gene clusters
in most cases and may be revised in future.

FEATURES
Location/Qualifiers
1. .210614
/organism="Streptomyces rochei"
/mol_type="genomic DNA"
/strain="7434AN4"
/db_xref="taxon:1928"
/plasmid="pSLA2-L"
/note="linear plasmid"
1. .1992
/note="left terminal inverted repeat, TIR-L; shows 99.4 %
(1981/1992) sequence identity to TIR-R (complement
(208623..210614))"
683..2188
/note="N-terminal sequence is almost identical (435/437)
with that of ORF143 at the right end of pSLA2-L until the
inner end of TIR-L
ORF1 (501 aa)
similar to AL590463 Streptomyces coelicolor putative
helicase, SCPI_136 (879 aa); homology is seen until the
inner end of TIR-L"
/codon_start=1
/transl_table=11
/product="putative helicase"
/protein_id="BAC76459.1"
/db_xref="GI:30698346"
/translation="MSTSRDQREAAQAEVDVAVRALELPVRLAPRGLRTQVIM
ATGSGKTRVAARSKEKLRGRVILVPSLDLITQTEAAWREAGTGMIGVSLRGED
VAFPNITDVEELVDWVRPFDKTVFATYASLGLELRHAGGLPGWDLIVVDEAHT
SGRLGKPAVVDHNTRIPSRLRYMTATPRMLQDLDEACGACELVSMEDPDGLFG
ARCFTLTSEAIDRGICAPYQVVCVDITDQLQAQLGVEGSDVRGRLAALQTA
LLKASSENFRFTLVPHMVKAEAPFAGLPDVAKLHAAGPLVPRTTWAWLCSGH
KFGHRRRLVGETSGIATDGTVEKGLGSKVLGEGVDTRECDSVYADVRGSMFPL

VOAQRALRIPQGGKVASLVVPLDPEETADNMLTGRPYNGLALLFRQAPLLTGHG
 EGFPAARPPGVEVPLVGRGAORAGSAGPASRSITPSAMSSLSISWRAP"
 complement (2315..3595)
 /note="ORF2 (426 aa)
 similar to AE004736-10 Pseudomonas aeruginosa hypothetical
 protein (442 aa)"
 /codon_start=1
 /transl_table=11
 /product="putative secreted protein"
 /protein_id="BAC76460.1"
 /db_xref="GI:30698347"
 /translation="MRKRALLGLIPLGLLVAVLAPAAHAASADTYVGTWAPPTAA
 PAADSTVEDQTLRQTHLSVAGDSLRVTFNEFGTSLTIGEVHAARPAAGPATAV
 DPGTVRFGGRPATAPGTQRMWDVALPTAGGLVLSLYLPQTGTSVHSA
 YGHNVAAGDVTGADPLTPVSTATSMHFLSGVSDRAGLADSVVTLGDSYTDGEHT
 TLDAERPDLLAERLRDGGLAGTGVNAGIGGNLLARDPDPGSAEASFAAYEGE
 SALKFRDRLQPGARAVTVLLGVNDLQPGGIAAPASDEVTAABELIAGIYLRSEAH
 EHLKLTATITFFAGDTTGYTFPREAVRQVNDWIRTSGAFDITVDFDAVLRDPAR
 PDHLLPAYDGGGLHNDAGMAARAFDLSDR"
 complement (3659..4645)
 /note="ORF3 (328 aa)
 similar to Y00459-2 Streptomyces griseus regulatory
 protein, StrR (350 aa)"
 /codon_start=1
 /transl_table=11
 /product="StrR-like regulatory protein"
 /protein_id="BAC76461.1"
 /db_xref="GI:30698348"
 /translation="MTAALAGNLDQETRIPISSLVPGYSPRLCGEDAEHVARLADTD
 EALPFLVERSLRVDIGMHRVLAKAGKHTIIEVLFDDGAEEAFLLAVRNMWTHGL
 PLRQDRRAAQRILIAQPHLSDRVAGIAGAKTVAALRPLAAGPTNPQARGRD
 GRIRLDGTTGRKAAELAQPOASREVRAHAGISATASDVRRLASGRSPYPER
 NTPGARPGCTGSRTAPGGGAPEPVSVRPPEDSPVRLRLDPSLRHESGRRL
 LLQCGAVERTALLAMACTVPPHCTDLVAELAREYADLWAEFAREVARTDG"
 complement (4922..5812)
 /note="ORF4 (296 aa), lankacidin biosynthesis protein
 similar to AE004625-4 Pseudomonas aeruginosa
 pyrroloquinoline quinone biosynthesis protein B (304 aa)"
 /codon_start=1
 /transl_table=11
 /product="pyrroloquinoline quinone biosynthesis protein B"
 /protein_id="BAC76462.1"
 /db_xref="GI:30698349"
 /translation="MILITGAAGGGFPPWNCALCARGRRGELPARSOECVAVSGDG
 RWMLNAGSDRTQLLAAPALTDPGPRDTPVRGLLTDAEVDHALGLAVLRGATKL
 TVAAPVPVAGSALPELVRLDRDWDATAPGGFAVAGGLTVTAHPVGTKAPK
 YAHAPDPAPWCAYRIEDPATGGALVYAPCLATWPDGFDLLASATCALLDGTFFSA
 GELGTATSAAGAGQSLMGLHPVAGPGGSLAALRHRLRIYTHLNTNPLDPPSAA
 HAARVAGVEVLPDGSSELV"
 complement (5815..6909)
 /note="ORF5 (364 aa), lankacidin biosynthesis protein
 similar to AJ277117-5 Gluconobacter oxydans putative pqgE
 protein (359 aa)"
 /codon_start=1
 /transl_table=11
 /product="pyrroloquinoline quinone biosynthesis protein E"
 /protein_id="BAC76463.1"
 /db_xref="GI:30698350"
 /translation="MADPAVGAAGMLIELTHRCPLHCPVCSNPLVREAEITCEQ
 WDTLQARELGVMVHFGSGELFAPDLPDLVGHARRLGAIVNLVTSQGLTAREAH
 DLARRGVHVQVLSQADPAAGQAIAGARVHTAKLEAARAVTAAGLPLTVNLVLRGN
 IDGTGRNLDVLDGADRIELANTQYIGLNRNLAALMPTAQLAAREAVRHARTYR
 AGPELVYAADYDRPKPCMDWGSTGLTVPAGDVLVPCPAAYAITLPPVENALRR
 PLSEIWAYSRSFNAYRGTMREPCRTCEPRHADHGGCRQAFQLTGDAAATDPACGL
 SPHRSVDAALAEVTDGVPVAFVPRGPVPA"
 complement (6902..7174)
 /note="ORF6 (90 aa), lankacidin biosynthesis protein
 similar to AL603642-200 Sinorhizobium meliloti putative
 pyrroloquinoline quinone biosynthesis protein D (98 aa)"
 /codon_start=1
 /transl_table=11
 /product="pyrroloquinoline quinone biosynthesis protein D"
 /protein_id="BAC76464.1"

/db_xref="GI:30698351"
 /translation="MTGLPEPTVRLRPGVRLTRDPARGELALLPERVVVNDTAAAV
 LAHCDGTTSLAGIVERLAEYEGVSAEDVREILLRLAQRVVVDLHG"
 complement (7171..7899)
 /note="ORF7 (242 aa), lankacidin biosynthesis protein
 similar to AE004625-5 Pseudomonas aeruginosa
 pyrroloquinoline quinone biosynthesis protein C (250 aa)"
 /codon_start=1
 /transl_table=11
 /product="pyrroloquinoline quinone biosynthesis protein C"
 /protein_id="BAC76465.1"
 /db_xref="GI:30698352"
 /translation="MSMSVTREVAAPWSEAFRQRLHALESYWDNRHFFRHRHGLL
 DEGLERLWAAWRYRCLPQKQRAIVANCLPEVRQWLRIIVVHDGADACAGAEK
 WRLRLAARVGLARDEVDRLVLAGTRFAVDVDFARRRPMLEAAASGLTELFSPGLL
 AHRGLRLEHVPWIAEYEGVEYFARIYVVGPEGRSLDLVARHAVRSREQEACVRALA
 PKCRVNAVLSLDYHTNGATRS"
 complement (7926..8030)
 /note="ORF8 (34 aa), lankacidin biosynthesis protein
 similar to AL603642-197 Sinorhizobium meliloti putative
 pyrroloquinoline quinone biosynthesis protein A (31 aa)"
 /codon_start=1
 /transl_table=11
 /product="pyrroloquinoline quinone biosynthesis protein A"
 /protein_id="BAC76466.1"
 /db_xref="GI:30698353"
 /translation="MRTSGKELPAKAWHRPDPFVITDTCMEVTAYFSR"
 8420..10057
 /note="ORF9 (545 aa), possible lankacidin resistance
 protein"
 Query Match 30.2%; Score 211.4; DB 1; Length 210614;
 Best Local Similarity 59.9%; Pred. No. 3.5e-17;
 Matches 391; Conservative 0; Mismatches 256; Indels 6; Gaps 2;
 QY 34 CAGGACGGCGCTCGGACGGCGGAGCGATCGTGGCGGAGCGCGCTCGTCTTCGAC 93
 DB 143115 CAGGACGGCGGATCGGACCGCGCGCGGTTCGAAAGCGGCGACCGATTTCGCC 143056
 QY 94 GAGTACGGGTTCGAGCGCGCCACAGTGGCAGAGATCTCTCGCGGCGCTCGTTCACCAAG 153
 DB 143055 GAGCAGGGTACCGCGCGCGACCGTGGCGGACATCTCAAGTGGCGCGCTGACCAAG 142996
 QY 154 GCGCGATGATCTTCATCTTCCTCCAAAGAAAGAGTGGCGCGCGCGCTGTCGGCCGAG 213
 DB 142995 GCGCGCTGATCTTCATCTTCCTCCAAAGAAAGCGCTTCGCCGCGGACCTCGGAGGC 142936
 QY 214 CAGACCTCGACGTGGCGGTGCGGAAATCCGGCTCCAAAGCGCGAGAACTGGTAGACCTC 273
 DB 142935 CAGGTGCGCGACAGCTGCTTCGCA---GCGAGCTCAAGCTCCAGGAATGGGTGGACGG 142879
 QY 274 ACCATGCTGGTCCCGCACCGCATGTGACAGATCCGATCCTGCGGCGCGGCGACCGCGCTC 333
 DB 142878 GGCATGACGTGGCGCACCGATGCGCGGAGCCCGTCTGCGCGCGCGCGCGCGCTC 142819
 QY 334 GCATGACAGCGGGCGGTGAGATCTTCGAGCGCAACCCGCTGCGCGAGTGGGGCGAC 393
 DB 142818 TCGCGCGAGCA---CACCGGAGCGAGCAGCAGCGCGCTTCGCCACCTGGATGCGC 142762
 QY 394 ATCTGCGCCAGCTCTGCGCGGAGGACAGGACAGGAGCGGGGAGGTGCTTCGCCACGTGAAC 453
 DB 142761 TTCTCGGCTCGCTCTTTCAGCAGGCGAAGCGGAGGAGGTCTTGGGCCACATCGAG 142702
 QY 454 CCGAAAAAGACCGGGGACATTCATGCTGCGGTCTTCCCGGGTCCAGCGGGTCTCCCGG 513
 DB 142701 CCGCGCGAGACCGCGAATGCGTCTCTCGGCTCTTCCAGCGGATCCAGCTCTCTCCAG 142642
 QY 514 GTCACTCCGACCGCGACCTCGCGACCGGATCTCGGTGATGGAACACACGTGCTG 573
 DB 142641 TTGACAGCAATGGCGGACATCGAGCAGCGTCTCGCGCTGTTCCGGCACGCTCTC 142582
 QY 574 CCAGCATGCTGCGCGCGTCCATGCTGACCTGGATCGAAACCGCGGAGGACGAGTCGGG 633

[illegible]

Db	628	CAGACCTCCACGGTGCAGTTTCGACGACGAGGGGTCCCGCTTCGACGTCCCTGGTGGACGGT	687
Qy	274	ACCATGCTGTCGCCACCGCATCTGTCACGATCCGATCCTGCGGGCGGGCACCGCGCTC	333
Db	688	GGCAGAGTTTCGTTTCGCTGTCGCCCAAACTCGATGGCGGGCCGGCACCGAGCTC	747
Qy	334	GCACTGCACGAGGGCGGTGGATTCCTCCGACGCCAACCCGTTCCGGCGAGTGGGGCGAC	393
Db	748	TCCATCGA-----GGGTGTCTTCCTCGCGGACCGCATCCATGGGCGACTGGATCGAC	801
Qy	394	ATTCGGCCGACGCTCTTCGGGAGGACAGGAAACGGGGGAGGTGCTTCGACACGTGAAC	453
Db	802	CGACGCGCCGGATGTCGAGCTGGGCCAGGACGCGCGGAGGTGTTCACACAGATCGAC	861
Qy	454	CCGAAAAAGACCGGCACATTCATCTGTCGGCTGCTTCACCGGGTCCACGCGGTCTCCCGG	513
Db	862	CCGATGGTGCACCAAAATCATCTGCTGCTGTTGTTCCACCGGTATTCAGCTGGTCTCGGAG	921
Qy	514	GTCACTCCGACCGCCAGGACCTCGGCCACCGGATCTCGGTGATGTGGAACACGTCGTCG	573
Db	922	GCCGATTCGGTGGGCCGATCTGCGGGGAGGTGGCGGAGATGTGGCCACATTCG	981
Qy	574	CCGAGATGTCGGCGGTTCATCTGACCTGATCGAAAACCGGCGAGGAGCGGATCGGG	633
Db	982	CCGTCGATCGCTCACCCCGCGTCATTGCGCCATCAAGCCGAGGCGCGTAGATCTG	1041
Qy	634	AAGTTCGGCGCGCGCGAGGCGCGGAGGCTCGGAGCGCTCGGAGCGCTCCGAGCGCGCTCCGAC	693
Db	1042	GCGGCCAGCGCGCGAGAGGCGGAGCGTGTAGGACAGAGGCCAGGTCGCCCGGAG	1101
Qy	694	GAGTAG 699	
Db	1102	GCACAG 1107	
RESULT 10			
LOCUS	AB121071		
DEFINITION	Kitasatospora setae gene for Ksba, partial cds.		
ACCESSION	AB121071		
VERSION	AB121071.1	GI:47550550	
KEYWORDS			
SOURCE	Kitasatospora setae		
ORGANISM	Kitasatospora setae		
REFERENCE			
AUTHORS	Choi,S.U., Lee,C.K., Hwang,Y.I., Kinoshita,H. and Nihira,T.		
TITLE	Cloning and Functional Analysis by Gene Disruption of a Gene Encoding a [gamma]-Butyrolactone Autoregulator Receptor from Kitasatospora setae		
JOURNAL	J. Bacteriol. 186 (11), 3423-3430 (2004)		
PUBMED	15150228		
REFERENCE	2 (bases 1 to 654)		
AUTHORS	Choi,S., Lee,C., Hwang,Y., Kinoshita,H. and Nihira,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (26-SEP-2003) Takuya Nihira, Osaka University, International Center for Biotechnology; 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan (E-mail:nihira@icb.osaka-u.ac.jp). Tel:81-6-6879-7452, Fax:81-6-6879-7454)		
FEATURES			
source	1..654		
	/organism="Kitasatospora setae"		
	/mol_type="genomic DNA"		
	/strain="IFO14216"		
	/db_xref="taxon:2066"		
	1..654		
	/gene="ksba"		
gene	<1..654		
CDS	/gene="ksba"		
	/codon_start=1		
	/transl_table=11		

ORIGIN

Query Match 28.3%; Score 197.8; DB 1; Length 654;
Best Local Similarity 59.2%; Pred. No. 7.7e-15;
Matches 337; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

/product="Ksba"
/protein_id="BAD20229.1"
/db_xref="GI:47550551"
/translation="MAESPRAAKTREAIIOAAGVFEQHGYNAGLSMTSRSLTGMG
AIYFHTSKDDLARAVMAQADLLVDPGERGLRLIDITLYLASELRQNLVLRAGVR
LAIEQSGFSDSTYLVQWDFRDELGAABVGLLPEVVADVLAQLLVSSYGTQL
LSEIATRADLPERIVRMWYLLPGIATPEARTLRLEPSVADRADDRNDQFGR"

QY 40 CGGGCGCTCGGCACGGCGGCGATCGTGGCGGACCGCTCGTCTTCGACGAGTAC 99
DB 16 CGGGCGGCAAGACCGCTGAGCGGATCATCCAGGCGCGCGCGGTGTTCGACGAC 75
QY 100 GGGTTCAGGCGGCCACAGTGCAGAGATCCTCTCGCGGGCTCGGTGCACCAAGGCGCG 159
DB 76 GGGTACAACGGCGCGCGGTGTGATGATCACTCCAGGTGGGGCTGACCATGGCGCG 135
QY 160 ATGTACTTCCACTTCGCTTCCAGGAGAGCTGGCGCGCGGTGCTGCGCGACGAC 219
DB 136 ATCTACTTCCACTTCACTTCCAGGACGACCTCGCGCGCGCGGTGATGGCGCGCG 195
QY 220 CTGCAGTGGCGGTGGCGGAATCCGGCTCCAAGGCGCAGGAACCTGGTAGACCTCA 279
DB 196 GAGGACCTGCTGCTGCCGACGGCGGCGGCTGCACGCGCTGATCGACATCACCTG 255
QY 280 CTGGTCCGCCACGGCATGCTGCACGATCCGATCTCTCGCGCGGCGACCGCGTCCG 339
DB 256 TACCTGGCTCGGAGTTGAGCGGAACGCTGCTGCGCGCGCGGCTCGCGTGGCGATA 315
QY 340 GACGAGGGCGGTGAGCTTCTCCGACGCAACCGCTTGGCGGAGTGGGGCGACATCT 399
DB 316 GAGCAGGGCTCGTTCGGCTCGCGGACAGACACCCCTACCTCGAGTGGGTGCGAC 375
QY 400 GCCAGCTCTCGCGGAGGACACAGGAACGGGGGAGGTGCTTCCGACGCTGAACCCG 459
DB 376 CGGGAGAGCTGGCGCGCGCGGAGCGCTGCGGCGACTGCTCCGAGGTGCTGCTGG 435
QY 460 AAGACGGGAGCTTATCTGCTGGCTGCTTCAACGGGCTCCAGCGGTCTCCGGGTAC 519
DB 436 GACGTCCGCCAACTGCTGGTGAAGTCTCTACAGCGGACGAGTGTGTCGAGATCG 495
QY 520 TCGACCGCGGACCTCGCGCACCGGATCTCGGTGATGTGGAACACAGTGTGCCAG 579
DB 496 ACCGACCGGCGGACCTGCGCGAGCGGATCGTCCGATGTGACCTTACCTGCTCC 555
QY 580 ATGTCGCGCGCTCCATGCTGACCTGGAT 608
DB 556 ATCGCCACCGGAGGCGCGCACCGGCT 584

RESULT 11
AB126048
LOCUS
DEFINITION
Kitasatospora setae 6997 bp DNA linear BCT 21-MAY-2004
Kitasatospora setae genes for homolog of putative metalloproteinase
protein, hypothetical protein, ksbA, BarB homolog, putative RNA
polymerase ECF-family sigma factor protein, homolog of
methylated-DNA-protein-cysteine methyltransferase protein, partial
and complete cds.
AB126048
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Kitasatospora setae
Kitasatospora setae
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycineae; Streptomycetaceae; Kitasatospora.
REFERENCE
1
AUTHORS
TITLE
Choi, S.U., Lee, C.K., Hwang, Y.I., Kinoshita, H. and Nihira, T.
Cloning and Functional Analysis by Gene Disruption of a Gene
Encoding a (Gamma)-Butyrolactone Autoregulator Receptor from

Kitasatospora setae
J. Bacteriol. 186 (11), 3423-3430 (2004)
15150228
2 (bases 1 to 6997)
Choi, S., Lee, C., Hwang, Y., Kinoshita, H. and Nihira, T.
Direct Submission
Submitted (11-NOV-2003) Takuya Nihira, International Center for
Biotechnology, Osaka University; 2-1 Yamadaoka, Suita, Osaka
565-0871, Japan [E-mail: nihira@icb.osaka-u.ac.jp,
Tel:81-6-6879-7452, Fax:81-6-6879-7454]
Location/Qualifiers
1..6997
/organism="Kitasatospora setae"
/mol_type="genomic DNA"
/strain="IFO 14216"
/db_xref="taxon:2066"
complement(1..1222)
/codon_start=1
/transl_table=11
/product="homolog of putative metalloproteinase protein"
/protein_id="BAD20230.1"
/db_xref="GI:47550553"
/translation="MSLLAVPAAASSGFHFGADGLDPYPTFGNGGVRVSHYDIRL
TYQATDLEGTTLLATTENLSRFNLDFAKVTEDVGRPAFAATGDOELSTTP
ARLPKGAETTVRYADVPSTVERYGTYTAWRTPDGALVAGEFAAAMWPPSNDHP
SDKATFDVSRVPAGTVISNGVLTGTSTRAGTWRDYDNRKPKQATYATLAIGLFDV
TQDTPGVPVNAVYSPDLGNAGARASVERTGBLVDWLSGYFGPYFSSVGLVFN
IPSSGGLTQTRIVYGPGLFARGSTSVVHELAHQWYGDVSUVAWMDJWLNEGFAQ
YAQFLWSHEGEGTALADMIYIYPADHPFWVTIPGDPDQGFADAVYQRGALAL
QALRTTVGDQVFLRI"
1548..2792
/note="homolog of Methanosarcina acetivorans C2A
hypothetical protein"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAD20231.1"
/db_xref="GI:47550554"
/translation="MTADGSDPRGAGLAPAAVATAVALGALVAVPGSAAHAQ
ERPGEEAHLVLTANTRANSVATYDPTASPVATVQVGSYPQAVSPDGSOAIVV
NNGSTVLTGTATGTAATVAGRLPGVAFSPDGAARYVGHVYSPRAGVDVDTA
TRTVATVPGSQPPGIATVPDGRVYVANLIGGSVIDAATSLTGTIVDSRRMP
AALAVPDGGRVYVANGSGLSVIDTASGAVSVITVAGGPGVSLSPDGAARYPS
QSDTIVSVDTGACRVATVPVGTPIGVLPADPAGAVYVANSKDGTLVLDPATDTV
SRTVPTGSSPQLATATVATVTVTVRAGHGPDPAGPPRPFEPDARGYAVVWKTDRAYAG
SCRVPALGLTGTVHTALFDFD"
3044..3583
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAD20232.1"
/db_xref="GI:47550555"
/translation="MTKTRTVGSGKPRPVRTDFVPCPSGEALSTRAAVCFGVGV
AATPLGSGTAGAAPPVPGADAVTVGPATVTVGPATVTVVCGRESSDTRAWNRPP
AKPAFRTVATVATVTLHPWRSIPRRPFGASAGASAGSGSVGRGAPGTDDGPEGR
LGMVGPVTASPSHAPLGR"
3904..4557
/gene="ksbA"
3904..4557
/gene="ksbA"
/codon_start=1
/transl_table=11
/product="ksbA"
/protein_id="BAD20233.1"
/db_xref="GI:47550556"
/translation="MAESPRAKTRERAIIOAAGVFEQHGYNAGLSMTSRSLTGMG
AIYFHTSKDDLARAVMAQADLLVDPGERGLRLIDITLYLASELRQNLVLRAGVR
LAIEQSGFSDSTYLVQWDFRDELGAABVGLLPEVVADVLAQLLVSSYGTQL
LSEIATRADLPERIVRMWYLLPGIATPEARTLRLEPSVADRADDRNDQFGR"
5288..5542
/codon_start=1
/transl_table=11
/product="BarB homolog"

/protein_id="BAD20234.1"
/db_xref="GI:4750557"
/translation="MVQRQAVPTGRLVLAASTAFDRAGFEGASPARTAOWAGVSGG
TUTVALVRSLAESAEVGAACGRPSRRSTRATPCRTATT"
5566. .6150
/codon_start=1
/transl_table=11
/product="putative RNA polymerase ECF-family sigma factor
protein"
/protein_id="BAD20235.1"
/db_xref="GI:4750558"
/translation="MDDVAEQLTELLAADDGFTLVVRHARVAHFLLRVSGSPAE
ADDLGDARLAVVALQGSPPERRELPRAMLMIAANWNRHVRGSRVVRVSPV
EBAATSTATGPARAEHGDRLVAALSELPHHRVPVVLHRHVLGLSVAEVAQV
QCPVTAQAQSVSGMDLRLGLAQEDLKEVTG"
6348. .6938
/codon_start=1
/transl_table=11
/product="homolog of methylated-DNA-protein-cysteine
methyltransferase protein"
/protein_id="BAD20236.1"
/db_xref="GI:4750559"

CDS

CDS

ORIGIN

Query Match 28.3%; Score 197.8; DB 1; Length 6997;
Best Local Similarity 59.2%; Pred. No. 4e-15;
Matches 337; Conservative 0; Mismatches 232; Indels 0; Gaps 0;
QY 40 CGGGCGCTCGACGCGGAGCGATCTGTCGGGCGAGCGCCCTCGTTCGACGAGTAC 99
DB 3919 CGCGCGCCCAAGACCGCTGAGCGCATCATCCAGCGCGCGCGCGTGTTCGAGCAGCAC 3978
QY 100 GGGTTCGAGCGCCACAGTGCAGAGATCCTCTCGGGGCGCTCGTTCACCAAGGCGCG 159
DB 3979 GGGTCAACGCGGCGCGGCTGTGATGATCACCTCAGTCGGGGCTGACCATGGCGGCC 4038
QY 160 ATGTACTTCACCTTCCTCCAGGAAGAGCTGGCGCGCGCGGTGTGCGCGGAGAGACC 219
DB 4039 ATCTACTTCCACTTCACTCCAGGACGACTCGCCCGGCGGTGATGCGCGCCAGGCG 4098
QY 220 CTGACGTGGCGTGCGGAAATCCCGGCTCAAGGCGCAGGAATCTGTAGACCTTCAACATG 279
DB 4099 GACGACCTGCTGTCGCCGAGCGGCGGCGCTGCAGCGCTGATCGACATCACCTTG 4158
QY 280 CTGGTCCGCAAGGATCTGCAGATCCGATCCTGCGGGCGGCGACGCGGCTCGCACTG 339
DB 4159 TACCTGGCTCCGAGTTGCAGCGGAACGTGCTGCTGCGGGCGGGGTCCGGTGGCGATA 4218
QY 340 GACGAGGGCGGTGACTTCTCCGACGCGCAACCCGTTGCGGAGTGCGGCGCACATCTGC 399
DB 4219 GAGCAGGGCTCGTTTCGGCGTCCGCGACAGCACCCCTACTGCAGTGGGTGACCGGTTT 4278
QY 400 GCCCAGCTCTGCGGAGGACAGGAAACGCGGGGAGGTGCTTCCGACGCTGAACCCGAAA 459
DB 4279 CGGGACGAGTGGGCGCGCGAGGCGGTGCGGCGACCTGCTCCCGAGGTGCTGCTGGCG 4338
QY 460 AAGACGGCGACTTCACTGCTGGCTCTTCAACGGGCTCCAGCGGCTCCCGGGTCAAC 519
DB 4339 GAGCTGCCCAACTGCTGTGCTAGCTCTTACAGCGGACACGAGCTGCTGTCCGAGATCGCC 4398
QY 520 TCCGACCGCAGGACCTCGGCGCAACCGGATCTCGGTGATGTGAACACAGCTGTGCCAGC 579
DB 4399 ACCGACCGGCGGACCTGCGCGAGCGGATCGTCCGATGTGACCTACCTGCTCCCGCGC 4458
QY 580 ATGTCGCCGCGCTCCATGTGACCTCGAT 608
DB 4459 ATCGCCACCCGCGAGGCCCGCACCCGGCT 4487

RESULT 12
AY256849
LOCUS

DEFINITION

Streptomyces ansochromogenes putative gamma-butyrolactone-binding
protein gene, complete cds.

ACCESSION
AY256849

VERSION
AY256849.1

KEYWORDS
GI:32765840

SOURCE
Streptomyces ansochromogenes

ORGANISM
Streptomyces ansochromogenes

REFERENCE
1 (bases 1 to 901)

AUTHORS
Li, W. and Tan, H.

TITLE
Sabr: an important gene involved in regulating nikkomycin
production in Streptomyces ansochromogenes

JOURNAL
Unpublished

REFERENCE
2 (bases 1 to 901)

AUTHORS
Li, W. and Tan, H.

TITLE
Direct Submission

JOURNAL
Submitted (16-MAR-2003) Molecular Microbiology, Zhong Guan Cun Bei
Yi Tiao No.13, Beijing 100080, China

FEATURES
Location/Qualifiers

source

1..901

/organism="Streptomyces ansochromogenes"

/mol_type="genomic DNA"

/db_xref="taxon:115647"

239..901

/codon_start=1

/transl_table=11

/product="putative gamma-butyrolactone-binding protein"

/protein_id="AAP87374.1"

/db_xref="GI:32765841"

/translation="NAQDDRAVTRRAVLRAAAVFAERGYAAATISEILKRAGVTKG
ALYFPGSKAALAQVLOBLTPEYHLRELKLOEWDAGMALARRLPPEFLLAGVR
ISADPRDVLGSAWPAWRLTSHALTEAKRGEVLPHVVPETAQVFLGAMVGAQFV
SQTLGAWDLDRTAALYGHLLAAIAAPVLTRLDTAPDRGARVTAARRSEDLSGI
AC"

CDS

ORIGIN

Query Match 28.0%; Score 196; DB 1; Length 901;
Best Local Similarity 57.9%; Pred. No. 1.2e-14;
Matches 368; Conservative 0; Mismatches 265; Indels 3; Gaps 1;
QY 16 GAACGGGTGGCAGTCGACAGAAACGGGCGCTCGCACGCGGCGGCGATCTCGCGGCA 75
DB 230 GGAGGGTGTGTGGCTCAGCAGACCGAGCGCTCGGACCGCGCGCGCTGCTCGCGGCC 289
QY 76 GCGGCTCGGTCTTCGACGAGTACGGGTTCGAGGCGGCGCACAGTGGCAGAGATCTCTCG 135
DB 290 GCGGACGCGGTCTTCGCGAGCGCGGATACGCGGCGCGCCACCATATCCGAGATCTCAAG 349
QY 136 GCGGCTCGGTCAACAGGGCGCGATGTACTTCCACTTCGCTTCCAAGGAAGAGCTGGCC 195
DB 350 GCGGCGGTGTCTCACCAAGGGCGCGCTGTACTTCCACTTCGACTTCCAAGGCGGCGCTCGCC 409
QY 196 GCGGCGGTCTGCGCGAGCAGACCTGTCAGCTGGGCGGTGCGGGAATCCGGCTTCAAGCGC 255
DB 410 CAAGGGTGTTCAGAGAACAGCTGACCGCCGAGTACCACTGCGCGGCGAGCTGAAGTTA 469
QY 256 CAGGAACCTGGTAGCTCACCATGCTGGTTCGCGCCACGGCATCTGCACCATCCGATCCCTG 315
DB 470 CAGGAATGGGTGGAACGCGGGGATGCGCTGCGGCGGCTGCCCGGGAACCTTCTCTG 529
QY 316 GCGGCGGGGCGCGCGGTCTGCACCTGGACACGAGGGGGGGTGGAGATTTCTCCACGCGCAACCGC 375
DB 530 CTCGCGGGGTCCGATCTCCCGGACCGCGCGCGCGCGTGAAGTGTCTCGGACGCGCTCG 589
QY 376 TTCGCGAGTGGGGGAGACATCTGCGCCAGCTCTTCTGGCGGAGGCGACAGGAAACGCGGGGAG 435
DB 590 CGGGCTGGGGCGGAGTACCTCGACGCGGCTCA---CGAGGCGCAAGAAACGCGGCGAG 646
QY 436 GTGCTTCGACAGCTGAACCCGAAAAGACCGGGGAGTTCATCGTGGCTGCTTTCACCGGG 495

```

Db      647  GTGTCGCCACAGTGTGTCGCCGAGAGACGGCCAGGTGTTCTCGGCGCTCGGTGGGC 706
QY      496  CTCACGGCGTCTCCGGGTACCTCCGACCCGACGACCTCGGCACCGGATCTCGGTG 555
Db      707  GCGCAGTTCGTCTCGACAGACCTCGCCGCTCGGCGGACCTGGACGACGACGCGCGCG 766
QY      556  ATGTGAACACAGTGTGTCGCCAGCATCGTCCGGGTCCATCTGACCTGGATCGAAACC 615
Db      767  CTGTAGGCACTCTCTCGCCGATCGGCGCCCGCCGCTCTCACCGGCTGACACC 826
QY      616  GCGCAGGACGATCGGGAAGTTCGCGCGCGGCC 651
Db      827  GCGCGGACCGCGCGCGGTGATCGCCGAGGCC 862

RESULT 13
AF156161
LOCUS      2637 bp      DNA      linear      BCT 10-NOV-1999
DEFINITION Streptomyces tendae TarA gene, complete cds.
ACCESSION AF156161
VERSION AF156161.1 GI:6318602
KEYWORDS Streptomyces tendae
SOURCE Streptomyces tendae
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
          Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE 1 (bases 1 to 2637)
AUTHORS Engel, P. and Scharfenstein, L.L.
TITLE Isolation of a gene specifying a gamma-butyrolactone-binding
        protein from Streptomyces tendae
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2637)
AUTHORS Engel, P. and Scharfenstein, L.L.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Food and Feed Safety, SRRC, USDA, 1100
        Robert B. Lee Blvd., New Orleans, LA 70124, USA
FEATURES
     Location/Qualifiers
     1..2637
        /organism="Streptomyces tendae"
        /mol_type="genomic DNA"
        /strain="ATCC31160"
        /db_xref="ATCC:31160"
        /db_xref="taxon:1932"
     626
        /note="transposon Tn4560 insertion site"
     1922..2584
        /function="nikkomycin biosynthesis"
        /note="gamma-butyrolactone-binding protein"
        /codon_start=1
        /transl_table=11
        /product="TarA"
        /protein_id="AAF06961.1"
        /db_xref="GI:6318603"
        /translation="MAQQDRAVTRRAVLRAAAVPAERGAAYAAIISILKRAQVTKG
        ALYHFDKAAALQVLSQTLPEYHLPRELQEWDMAGMTLARRLPPEPLLAGVR
        ISADRGVLSAWPAWRLTSHVLEAKRGVLPVHPVPEETAQVFLGAWGAQFV
        SOTLAWEDLDRLTAALYSHLLGATAAPPVLTLDTPDRGARVIAEARRRSGDLGI
        AC"

misc_feature
626
CDS
     1922..2584
        /function="nikkomycin biosynthesis"
        /note="gamma-butyrolactone-binding protein"
        /codon_start=1
        /transl_table=11
        /product="TarA"
        /protein_id="AAF06961.1"
        /db_xref="GI:6318603"
        /translation="MAQQDRAVTRRAVLRAAAVPAERGAAYAAIISILKRAQVTKG
        ALYHFDKAAALQVLSQTLPEYHLPRELQEWDMAGMTLARRLPPEPLLAGVR
        ISADRGVLSAWPAWRLTSHVLEAKRGVLPVHPVPEETAQVFLGAWGAQFV
        SOTLAWEDLDRLTAALYSHLLGATAAPPVLTLDTPDRGARVIAEARRRSGDLGI
        AC"

ORIGIN
Query Match 27.8%; Score 194.4; DB 1; Length 2637;
Best Local Similarity 57.7%; Pred. No. 1.3e-14;
Matches 367; Conservative 0; Mismatches 266; Indels 3; Gaps 1;

QY      16  GAACGGGTGCGACGACGAGGACGGCGTCCGACCGCGGCGGATCGTGGCGGCA 75
Db      1913  GGAGGTGTGTGGCTCAGCAGGACCGGACCGTCCGACCGCGCGCGTGTGCGCGGCC 1972
QY      76  GCGCGCTCGGTCTTCACACAGTAGTACGGGTTCGAGCGCGCCACACAGTGGCAGAGATCTCTCG 135
Db      1973  GCGGACGGGTCTTCGCGGAGCGCGGATACGCGCGCGCCACCATCTCCGAGATCTCTCAAG 2032

```

```

QY      136  CGGCGCTCGGTCAACAAGGGCGCGATGTACTTCACATTTCGTTTCAAGGAAGAGTGGCC 195
Db      2033  CGGCGCGGTGTCAACAAGGGCGCGCTGTACTTCACATTTCGACTTCAAGAGCGCGCTCGCC 2092
QY      196  CGGCGGTGTGCGGCGGACAGACCTGACCTGCGGGTGGCGAATCCGGCTTCAAGGCG 255
Db      2093  CAAGGGGTGTTGCGAGGAACAGCTGACACCCGAGTACCACTTCCGCGCGAGCTGAAGTTA 2152
QY      256  CAGGAACCTGTGTAGACCTCACCATCTGTGTGTCGCCCAACGGCATGTCTGCACGATCCGATCCTG 315
Db      2153  CAGGAGTGGGTGACGCGGGGATGACCTTGGCCAGACGGCTGCCCGGGAACCTTCTCTG 2212
QY      316  CGGCGGCGACCGCGCTCGCACTGACACGAGGGGGGTGGACTTCTTCGACGCGCAACCG 375
Db      2213  CTGCGCGCGCTCGGATCTCGCGGACCGCGCGCGGAGGTGCTCGGACGCGCTGG 2272
QY      376  TTCGCGAGTGGGCGACATCTGCGCCCGAGCTCTTGGCGGAGGACACAGAACGGGGGAG 435
Db      2273  CCGGCTGTGTACGGCTGACCT---CGCACGTGTCTACCGAGGCCAAGAGCGGGGCGAG 2329
QY      436  GTGCTTCGCGACGTGAACCCGAAAGAACCGCGCATTTTCATCGTCGGTGTTCACCGGG 495
Db      2330  GTGCTGCCGCGACGTGTGTCGCCGAGGAGACGCCAGGTGTTCTCTCGGCGCTGGGTGGC 2389
QY      496  CTCACGCGGTCTCCGGGTCACTCCGACCGCGCAGGACCTCGGCCACCGGATCTCGGTG 555
Db      2390  GCGCAGTTCGTCTCGCAGACCTCTCGCGCTGGGAAGACCTGGACGACCGCAGCGCGCG 2449
QY      556  ATGTGGAACACGTGTCGCCAGCATCGTCCGGGTCCATGCTGACCTGGATCGAAACC 615
Db      2450  CTGTACACCCACTCTCTCGGCGGATCGGCGCGCGCGCTCTCACCGGCTGACACC 2509
QY      616  GCGCAGGAGCGGATCGGGAAGTTCGCGCGCGGCC 651
Db      2510  GCGCGGACCGCGGTGCCGGGTGATCGCGAGGCC 2545

RESULT 14
AF117439/c
LOCUS      92294 bp      DNA      linear      BCT 01-FEB-2003
DEFINITION Streptomyces carzinostaticus subsp. neocarzinostaticus
          neocarzinostatin biosynthetic locus, complete sequence.
ACCESSION AF117439
VERSION AF117439.1 GI:28192466
KEYWORDS Streptomyces carzinostaticus subsp. neocarzinostaticus
          Streptomyces carzinostaticus subsp. neocarzinostaticus
          Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
          Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE 1 (bases 1 to 92294)
AUTHORS Liu, W., Nonaka, K., Nie, L., Bae, J., Zazopoulos, E., Farnet, C.M. and
        Shen, B.
TITLE Biosynthesis of the chromoprotein enediyne neocarzinostatin in
        Streptomyces carzinostaticus: analysis of the gene cluster and
        deduction of the biosynthetic pathway
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 92294)
AUTHORS Liu, W., Nonaka, K., Nie, L., Bae, J., Zazopoulos, E., Farnet, C.M. and
        Shen, B.
TITLE Direct Submission
JOURNAL Submitted (03-JUN-2002) Ecopia BioSciences Inc, 7290 Frederick
        Banting, Saint-Laurent, QC H4S 2A1, Canada
FEATURES
     Location/Qualifiers
     1..92294
        /organism="Streptomyces carzinostaticus subsp.
        neocarzinostaticus"
        /mol_type="genomic DNA"
        /strain="ATCC 15944"
        /sub_species="neocarzinostaticus"
        /db_xref="ATCC:15944"
        /db_xref="taxon:167636"
        10097..10894
        /codon_start=1

```

CDS


```

/protein_id="AAM77989.1"
/db_xref="GI:28192476"
/translation="MILAGKGTSLRPYTSRLPKPLVPIGDFEISLEIVLSQSRGP
DVTILAIGLHLIRSVFGGKRWNNHITVASEDTPRGITGVPLGILDOQLPHELFVNM
GDLTDLADVLAHTRGEAPLTVATCTREVRIDFGVLTTSNRVHFVEKPTVDYR
VSMGVGLSRITLLETTPSHMPEFGDLDLLARGVQPAEYFPDGYWLDIGRPDDYDR
ANAEFDSIKGTLLKGA"
complement (27203..28198)
CDS
/codon_start=1
/transl_table=11
/product="NDP-hexose-4,6-dehydratase"
/protein_id="AAM77990.1"
/db_xref="GI:28192477"
/translation="MAGLAVTGAEGFIGSHLVETLVREGRVRAVVOVNSFASWGL
DLDAEVMHVEFPGDVPDPAVLEVPASVWVHLLAIPYSYQAPRPSYLTNV
MGTNVLEAVRATGPRLVHTISSETISGTVSARVPIGDEHPLOQAPYAAKVAADLV
ESTHUSYGVAPVLPFRFNTGPRQSAVAIPTIIISQVAGARQITGLSAPRTDFTV
TDAEAFVSGTAPASAVIGFNAGTGTVEIGRLAADIMRLMDADASIRDPQRLR
PKDSEVLRLVCDAGSLRARTGWRPRTDREDGLRRTIEWFRNPVQAGYRQTHYR"
complement (28341..29060)
CDS
Query Match      25.9%; Score 180.8; DB 1; Length 92294;
Best Local Similarity 56.8%; Pred. No. 2.1e-13;
Matches 332; Conservative 0; Mismatches 252; Indels 0; Gaps 0;
QY 21 GGTGGCAGTGGCAGCAGAACGGGCGTCCGCACGGCGGAGCGATCGTCGGGCGACCGC 80
Db 69432 GCCGGAACCGAGCAGACGGGCCACAGGACACAGAGCTTGATCTCGTGCCGCTGC 69373
QY 81 CTCGGTCTTCGACGAGTACGGGTTTCGAGCGCGCCACAGTGGCAGAGATCCTCTCGCGGCG 140
Db 69372 TGAGGTGTTTCGACGAGAGGGATACAGCGGGCGAGCATCAGTAAATCATGAACCGCGC 69313
QY 141 CTCGGTCCACAGGGCGGATGACTTCCACTTCGCTTCCAGGAGAGTGGCGCGCG 200
Db 69312 CGCGGTGACCCAGGGGGGATGACTTCCACTTCAAGTCGAAGAGAGTGGCGCGTGGC 69253
QY 201 CGTGTCTGGCGGACGACCTCGACCTGGCGGTTCGCGGAATCGGTCCTCAAGCGCAGGA 260
Db 69252 AGTGATGACAGTCACGACGACCTTCATCGACTTCCCTCAGAGGTGCGGCTGCAGCG 69193
QY 261 ACTGTTAGACTCACCATGCTGTGTCGCCACGCGATGCTGCGACGATCGATCTCGCGGC 320
Db 69192 AGTGATCGACCTCACCTTCACCTCGCTCACGAGTCCAGACCAACCCCTGATCGGCG 69133
QY 321 GGGCAGCGGCTCGCACTCGACAGGGGGGGTGGACTTCTCGAGCGCAACCCGTTGG 380
Db 69132 CAGCATCCGGCTCGCGTTCGAGCAGGAGAGTTCGGCGTTCGCGACGACACCGCCCTACCA 69073
QY 381 CGAGTGGGCGACATCTGCGGCCAGCTCTGCGCGAGGACAGAACGGGGGAGGTGCT 440
Db 69072 GGAATGGGTAGCCCTGGTGGAGACTTACCTGCTGGAGGCGGCCAGCGGCGACTGCT 69013
QY 441 TCGCAGCTGAACCCGAAAGACCGGCGACTTCATCTGTCGGGTGCTTCAACGGGCTCA 500
Db 69012 CCGGAGGTGGACGAACACGACTTCGCCGCTGCTTCGTCAGCGCGTTCACCGGTAGCA 68953
QY 501 GCGGGTCTCCGGGTTCACCTCGACCGCGCAGACCTTCGCGCACCGGATCTCGGTATG 560
Db 68952 GCTGTTCTGAGGCTCGCCAGGCGCGCTCGGATCTGCCAGCAGGATCTGCTGCTG 68893
QY 561 GAACCACTGCTGCCAGCATCTGTCGGCGCTGCCATGCTGACT 604
Db 68892 GCGTTATCTGCTGCTGGCTCAGCGCCCTCGGACGAGGGCT 68849

```

RESULT 15
 ABI06894
 LOCUS
 DEFINITION
 Streptomyces clavuligerus scar gene for butyrolactone autoregulator
 receptor protein, complete cds.
 ACCESSION
 ABI06894
 VERSION
 ABI06894.1 GI:29335760

```

Streptomyces clavuligerus
Streptomyces clavuligerus
Bacteria; Actinobacteridae; Actinobacteriales;
Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE
1
Kim,H.S., Lee,Y.J., Lee,C.K., Choi,S.U., Yeo,S.H., Hwang,Y.I.,
Yu,T.S., Kinoshita,H. and Nihira,T.
Cloning and characterization of a gene encoding the
gamma-butyrolactone autoregulator receptor from Streptomyces
clavuligerus
Arch. Microbiol. 182 (1), 44-50 (2004)
JOURNAL
PUBMED
15257430
REFERENCE
2 (bases 1 to 1101)
Kim,H.S., Lee,Y.J., Lee,C.K., Choi,S.W., Yeo,S.H., Hwang,Y.I.,
Yu,T.S. and Nihira,T.
Direct Submission
Submitted (26-MAR-2003) Hyun Soo Kim, Keimyung University,
Microbiology; 1000 Shindang-dong, Daegu 704-701, Korea
(E-mail:hakim@kmu.ac.kr. Tel:82-53-580-5284, Fax:82-53-580-6447)
FEATURES
Location/Qualifiers
1..1101
/organism="Streptomyces clavuligerus"
/mol_type="genomic DNA"
/db_xref="taxon:1901"
202..900
/gene="scar"
202..900
/gene="scar"
/transl_table=11
/product="butyrolactone autoregulator receptor protein"
/protein_id="BAC66444.1"
/db_xref="GI:29335761"
/translation="MARQRAIRTRKILVVAASVDFQVGEATISDILDSQSGTKG
ALYHFGSKELAQALAEVAPLPVPPQGLKEAIDEGMLLAHLITGEHDPDIIQ
GSIRLTVDQSSKDNLDKRVPMQWIDHSLGLFSARKNGLVPHADLDSVSLFAG
VTGVQLSRIMTDRLDIAERMSDMYRNLMPALAVPSVLARLDFSTRDALVYVRVMA
SQEQDEGAGYTVGV"
Query Match      24.8%; Score 173.4; DB 1; Length 1101;
Best Local Similarity 56.8%; Pred. No. 5.6e-12;
Matches 387; Conservative 0; Mismatches 276; Indels 18; Gaps 3;
QY 28 GTGCACAGGAAACGGGCGTCCGACGGGAGGATCGTCGGGCGACCGCCCTCGGTC 87
Db 205 GCGCGCAGGAGCAGGACCATCGGACGCGGAGAGATCCTGTGCTGCCACCGTC 264
QY 88 TTGACAGGATACGGGTTTCAGGCGCCACAGTGGCAGAGATCCTCTCGCGGCTCGGTC 147
Db 265 TTCGACAGGTGGGTCTACGAGCGCGGACCATCTCGACATCTCGACAGTCGGGGCTG 324
QY 148 ACCAAGGGCGGATGATTCACATTCGTTCCAAAGGAAGAGCTGGCGCGCGGCTGCTG 207
Db 325 ACCAAGGGGGGCTCTACTTCCACTTCGAGTTCGAAGGAGGAGTTGGCCCAAGCGTCTC 384
QY 208 GCGGAGCAGACCTCGACGT---GGCGGTGCGGGAATCGGCTCCAAGCGCAGGAAC 264
Db 385 GCGGAGCAGGTGCGCGGCTTGGCGCGGTCTCTCGCAGGGGCTCAAGTTCAGAGGGG 444
QY 265 GTAGACCTCACCATGCTGGTGGCCCA-----CGGCATGCTGCACGATCCGATCCTG 315
Db 445 ATCGATAGGGCATGCTCTCGCCCATCTCTCACCGGGAGCAGGTCAGCCGATCATC 504
QY 316 CGGCGGGGCGACCGGCTCGACTGCAACGAGGGGG-----GGTGGACTTCTCCACGCC 369
Db 505 CAGGGGAGCATCCGGCTCACCGTGACACGAGGCTCTCTCAAGGACAAATCTTCGACCGCG 564
QY 370 AACCGGTTGGGAGTGGGGGACATCTGCGCCAGCTCTCGCGGAGGACGACGAAACGG 429
Db 565 GTGCCCATGACAGGCTGGATCGATCAGCCTCGGGCTGTTCAGCGAGGCGCGGAGAAC 624

```

KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS

TITLE

JOURNAL
 PUBMED

REFERENCE
 AUTHORS

TITLE
 JOURNAL

FEATURES
 source

gene

CDS

ORIGIN

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2005, 19:44:49 ; Search time 588 Seconds
(without alignments)

7037.244 Million cell updates/sec

Title: US-10-049-710A-1

Perfect score: 699

Sequence: 1 atggcagtcgacacgaacg.....agggccgctccgacgagtag 699

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	699	100.0	699	6	AAL44719
2	229	32.8	4346	6	ABN86379
3	229	32.8	4346	10	ADSL4794
c 4	126	18.0	12905	3	AZ99101
5	126	18.0	12905	3	AZ99100
c 6	88.6	12.7	1819	8	ABV75877
7	85.8	12.3	9521	4	AAS07627
8	78.2	11.2	4257	2	AAV68520
9	78.2	11.2	4257	2	AAV10362
10	68.2	9.8	1224	8	ACA37971
c 11	68.2	9.8	32329	12	AD051695
c 12	67.8	9.7	12001	2	AAQ76213
c 13	67.2	9.6	1041	8	ACA42557
14	67.2	9.6	1041	8	ACA42557
15	66	9.4	1107	11	ABD05332
16	66	9.4	1134	11	ABD05304
c 17	66	9.4	1326	11	ABD05398
18	65	9.3	77536	3	AAI4651
c 19	64	9.2	9521	4	AAS07627
20	63.4	9.1	3354	8	AAL61181

21	63.4	9.1	5862	10	AAD54224	Aad54224 Streptomy
22	63.4	9.1	52101	10	AAD54217	Aad54217 Streptomy
c 23	63.4	9.1	82746	8	AAL61224	Aal61224 Actinosyn
24	63.2	9.0	1227	10	ADC36139	Adc36139 Weed cont
25	63.2	9.0	1439	10	ADC36142	Adc36142 Weed cont
26	62.8	9.0	110000	4	AAI99682_36	Continuation (37 o
27	62.8	9.0	110000	4	AAI99683_36	Continuation (37 o
28	62.6	9.0	582	8	ACA23667	ACA23667 Prokaryot
29	62.4	8.9	819	8	ACA24088	ACA24088 Prokaryot
30	62.4	8.9	4770	8	AAL61172	Aal61172 Actinosyn
c 31	62.4	8.9	109519	5	AAS08693	Aas08693 Micromono
32	61.8	8.8	1134	12	ADJ35079	Adj35079 DNA encod
33	60.8	8.7	5877	6	ABS78701	AbS78701 S. citric
34	60.6	8.7	381	13	ACN62234	ACn62234 Cotton gy
35	60.2	8.6	23673	6	ABZ75344	Abz75344 Human R11
36	60.2	8.6	27541	4	AAD17185	Aad17185 Streptomy
37	60.2	8.6	125401	4	AAD17186	Aad17186 Streptomy
c 38	60	8.6	27541	4	AAD17185	Aad17185 Streptomy
c 39	60	8.6	125401	4	AAD17186	Aad17186 Streptomy
c 40	59.6	8.5	2721	2	AAQ10543	Aaq10543 BamHI J-I
c 41	59.6	8.5	2721	2	AAQ10212	Aaq10212 BamHI J-I
42	58.8	8.4	9975	8	AAL61173	Aal61173 Actinosyn
43	58.4	8.4	642	8	ABZ71163	Abz71163 S. muraya
c 44	58.4	8.4	36321	8	ABZ71131	Abz71131 Streptomy
45	58.2	8.3	1386	13	ADS56031	AdS56031 Bacterial

ALIGNMENTS

RESULT 1
AAL44719
ID AAL44719 standard; DNA; 699 BP.
XX
AC AAL44719;
XX
DT 03-MAY-2002 (first entry)
XX
DE S virginiae butanolide binding protein coding sequence.
XX
KW Butanolide binding protein; gene expression induction; operator;
KW transgenic plant; antibiotic production; gene; ds.
XX
OS Streptomyces virginiae.
FH Key Location/Qualifiers
FT CDS 1..699
FT /tags= a
FT /product= "butanolide binding protein"
XX
PN WO200196581-A1.
XX
PD 20-DEC-2001.
XX
PF 15-JUN-2001; 2001WO-JP005096.
XX
PR 15-JUN-2000; 2000JP-00180466.
PA (KANF) KANEKA CORP.
XX
PI Shimmyo A, Kato K, Yamada Y, Nihira T, Shindo T;
XX
WPI; 2002-098073/13.
DR P-PSDB; AAM48990.
XX
PT Inducing expression of gene under regulation by operator at actinomycetes
PT self-regulator provision site, useful in producing transformant tobacco
PS for production of antibiotics e.g. virginiamycin.
XX
PS Example 1; Page 46-48; 57pp; Japanese.
XX
CC The present invention relates to a method of inducing the expression of a
CC gene under the regulation of an operator at an actinomycetes self-

CC regulator provision site comprising imparting the characters of a
CC repressor and an operator constituting a gene expression-inducing system
CC with use of the actinomycetes self-regulator as inducer to a plant by
CC gene transfer thereby providing the self-regulator to be transformed
CC plant. The method is useful in producing transformant tobacco for
CC production of antibiotics e.g. virginiamycin. The present sequence is the
CC Streptomyces virginiae butanolide binding protein coding sequence which
CC was used in the exemplification of the invention
XX
SQ Sequence 699 BP; 116 A; 233 C; 253 G; 97 T; 0 U; 0 Other;

Query Match 100.0%; Score 699; DB 6; Length 699;
Best Local Similarity 100.0%; Pred. No. 1.3e-110;
Matches 699; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCAGTCGACACAAACGGGTGGCAGTGCACAGAAACGGGCGCTCGCAGCGGGCAG 60
Db 1 ATGGCAGTCGACACAAACGGGTGGCAGTGCACAGAAACGGGCGCTCGCAGCGGGCAG 60
QY 61 GCGATCGTCGGGCGACGCCCTCGGTCTTCGACGAGTACGGTTCGAGGCCGCCACAGTG 120
Db 61 GCGATCGTCGGGCGACGCCCTCGGTCTTCGACGAGTACGGTTCGAGGCCGCCACAGTG 120
QY 121 GCAGAGATCCTCTCGCGGCGCTCGGTACCAAGGGCGGATGACTTCCACTTCGCTTC 180
Db 121 GCAGAGATCCTCTCGCGGCGCTCGGTACCAAGGGCGGATGACTTCCACTTCGCTTC 180
QY 181 AGGAAGACTGCGCCCGCGCTGTCGGCGAGCAGACCTCGACGTGGCGGTGCGGAA 240
Db 181 AGGAAGACTGCGCCCGCGCTGTCGGCGAGCAGACCTCGACGTGGCGGTGCGGAA 240
QY 241 TCGCGCTCAAGCGCAGGAACTGGTAGACCTCACCATCTGTCGCCACCGCATGCTG 300
Db 241 TCGCGCTCAAGCGCAGGAACTGGTAGACCTCACCATCTGTCGCCACCGCATGCTG 300
QY 301 CACGATCCGATCTCGCGGCGGCGACCGCGCTCGCACTGGAACAGGGCGGTGGAATTC 360
Db 301 CACGATCCGATCTCGCGGCGGCGACCGCGCTCGCACTGGAACAGGGCGGTGGAATTC 360
QY 361 TCGGAGCCAAACCGGTTCGGCGAGTGGGCGACATCTGCGCCAGCTCTCGCGGAGGCA 420
Db 361 TCGGAGCCAAACCGGTTCGGCGAGTGGGCGACATCTGCGCCAGCTCTCGCGGAGGCA 420
QY 421 CAGGAACCGGGGAGGTGCTTCGACAGTGAACCCGAAAGAACCGCGCACTTCATCGTC 480
Db 421 CAGGAACCGGGGAGGTGCTTCGACAGTGAACCCGAAAGAACCGCGCACTTCATCGTC 480
QY 481 GGTGCTTCAACGGGCTCCAGGGGCTCCCGGGTCACTTCGACCGCGGACCTTCGGC 540
Db 481 GGTGCTTCAACGGGCTCCAGGGGCTCCCGGGTCACTTCGACCGCGGACCTTCGGC 540
QY 541 CACCGATCTCGGTGATGTGGACACAGTGTGCGCCAGCATCGTGGCGGCTCCATGCTG 600
Db 541 CACCGATCTCGGTGATGTGGACACAGTGTGCGCCAGCATCGTGGCGGCTCCATGCTG 600
QY 601 ACTCGATCGAAACCGCGGAGAGCGGATCGGGAAGGTTCGGCGGCGCGCGAGCGGCC 660
Db 601 ACTCGATCGAAACCGCGGAGAGCGGATCGGGAAGGTTCGGCGGCGCGCGAGCGGCC 660
QY 661 GAGGCTCGGAGGCTTCGAGGCGGCTTCGACAGTAG 699
Db 661 GAGGCTCGGAGGCTTCGAGGCGGCTTCGACAGTAG 699

RESULT 2

ABN86379
ID ABN86379 standard; DNA; 4346 BP.
XX
AC ABN86379;
XX
AT 21-OCT-2002 (first entry)
DT
XX ScbA, ScbR and ScbB encoding genes containing DNA sequence.

XX Antibiotic; bacterium; scbA; afsA; scbR; arpA; barA; actinorhodine; Act;
KW undecylprodigiosin; Red; gene; ds.
XX Streptomyces coelicolor.
XX CA2322241-A1.
XX 23-APR-2002.
XX 23-OCT-2000; 2000CA-02322241.
XX 23-OCT-2000; 2000CA-02322241.
XX (PLAN-) PLANT BIOSCIENCE LTD.
XX Takano E, Bibb M;
XX WPI; 2002-501089/54.
XX P-PSDB; ABN80940, ABN80941, ABN80942.
XX Modifying antibiotic-producing Streptomyces, to increase, or alter timing
PT of, antibiotic production, by deleting the scbA or scbR genes.
XX Claim 19, 20; Fig 14; 64pp; English.
XX The invention provides a method for modifying an antibiotic-producing
CC strain of Streptomyces to increase production of antibiotics or to alter
CC the timing of antibiotic production. The modification is functional
CC deletion of the scbA gene of S. coelicolor, or its homologues, but is not
CC deletion of the afsA gene of S. griseus, or the modification is
CC functional deletion of the scbR gene of S. coelicolor, or its homologues,
CC but is not deletion of arpA of S. griseus nor barA of S. virginiae. The
CC method is particularly used for production of the antibiotics
CC actinorhodine (Act) and undecylprodigiosin (Red). The present sequence
CC represents the nucleic acid sequence of the region containing the S.
CC coelicolor ScbA, ScbR and ScbB proteins
XX
SQ Sequence 4346 BP; 650 A; 1561 C; 1519 G; 616 T; 0 U; 0 Other;

Query Match 32.8%; Score 229; DB 6; Length 4346;
Best Local Similarity 60.8%; Pred. No. 2.1e-30;
Matches 373; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 34 CAGGAACCGGGCGCTCGCACCGCGCAGCGATGTCGGGCGACCGCTTCGTTCGAC 93
Db 2270 CAGGACCGGGCGATCGCACCGCGCAGACGATCTTGACCGCGCGCGAGTCTTCGAG 2329
QY 94 GAGTACGGGTTCGAGCGGCCACAGTGGCAGAGATCTTCGCGGGCTTCGGTACCAAG 153
Db 2330 AAGCAGGGCTACCAAGTCCACGATCAGGAGATCTCAAGTGGCGGGTGACCAAG 2389
QY 154 GCGCGATGATCTTCACCTTCGCTTCCAGGAGAGCTGGCCCGCGGCTTCGCGCGAG 213
Db 2390 GGAGCGCTCTACTTCACCTTCAGTTCAGGAAGAACTGGCGCTGGCGCTTCGACGCC 2449
QY 214 CAGACCTTCAGCTGGCGGTGCGGAATCCGCTCCAGGCGCAGAACTGGTAGACCTC 273
Db 2450 CAGGAACCAACACAGCGCGCTTCGAGGCAACCCCTCCGCTCGAAGAACTCATCGACATG 2509
QY 274 ACCATGTGGTTCGCGCACGGCATGTGCAGATTCGATCTTCGCGGGCGGCGACGCGGCTC 333
Db 2510 GGCATGTTGTTCTGTACCGCTTGGCGCACGAACGTGTGGCGCGGCGCGGCTTCGCGCTC 2569
QY 334 GCACCTGGACACGCGGCGGTGGACTTTCGACGCGCAACCCGCTTCGCGGAGTGGGCGGAC 393
Db 2570 TCCATGGACACAGCGCACCGTCTCGATCCCGAGGACCCCTTCGCTCGCTGGCACGAG 2629
QY 394 ATCTGCGGCCAGCTCTTCGCGGAGGACAGGAACCGGGGAGGTGCTTCGCGACGTTGAC 453
Db 2630 ACATCTCTGAAGCTGCTGTAACCCAGGCGCAGGAGAAACGGTGAAGTGTGTCGCCCATGTGGTC 2689
QY 454 CCGAAAGACCGCGGAGCTTCATCTGTCGCTCTTCCACCGGCTCCAGCGGCTCTCCCGG 513

Db 2690 ACCACGATCGGCCGATCTTACGTGGGACGTTCCCGGGATACAGTGTGTCCTCCAG 2749
Qy 514 GTCACTCCGACCGCCAGGACCTCGGCCACCGGATCTCGGTGATGTGGAACCAACGCTGCTG 573
Db 2750 ACGGTACGCGACTTACGAGACCTCGAACCGCTACGCGTCTGTCGAGAAGCACATCTCTG 2809
Qy 574 CCCAGCATGTGCGCGGCTCCATGCTGACCTGATCGAAACCGGCGAGAGCGGATCGGG 633
Db 2810 CCGGCATCGCGTTCCTCCGCTGCTGCGCGCTGATCTCTCCGAGGAGCGCGAGCA 2869
Qy 634 AAGGTCCGCGCG 646
Db 2870 CGCCTCGCGCG 2882

RESULT 3

ADE14794

ID ADE14794 standard; DNA; 4346 BP.

XX ADE14794;

XX ADE14794;

XX 29-JAN-2004 (first entry)

XX Streptomyces coelicolor partial genome sequence. 2.

DE antibiotic-producing strain; antibiotic production; scbA gene; Scbr gene;

XX actinorhodin; undecylprodigiosin; gene; ds.

KW Streptomyces coelicolor.

XX

FH Location/Qualifiers

FT complement (1199..2143)

FT /tag= a

FT /product= "Streptomyces coelicolor Srbs protein"

FT /transl_except= (3149..3147, aa:Tyr)

FT 2261..2308

FT /tag= b

FT /product= "Streptomyces coelicolor Srbr protein"

FT complement (3024..3812)

FT /tag= c

FT /product= "Streptomyces coelicolor Srba protein"

FT /transl_except= (3150..3145, aa:Met)

FT

XX US2003124644-A1.

XX

XX 03-JUL-2003.

XX

XX 23-OCT-2001; 2001US-00017471.

XX

XX 23-OCT-2000; 2000US-0242561P.

XX

XX (TAKA/) TAKANO E.

XX (BIBB/) BIBB M J.

XX

XX Takano E, Bibb MJ;

XX

XX WPI; 2003-810983/76.

XX

XX P-PSDB; ADE14791, ADE14792, ADE14793.

XX

XX Modifying an antibiotic-producing strain of Streptomyces coelicolor or

XX Streptomyces lividans to increase or alter the timing of antibiotic

XX production in the strain, comprises functionally deleting in the strain

XX the scbA or Scbr gene.

XX

XX Claim 19; Fig 14; 33pp; English.

XX

XX This invention relates to the novel modification of an antibiotic-

XX producing strain of Streptomyces coelicolor or Streptomyces lividans to

XX increase or alter the timing of antibiotic production in the strain.

XX The method comprises functionally deleting in the strain the scbA or Scbr

XX gene. The method is useful in increasing and altering the timing of

XX antibiotic production (especially actinorhodin and undecylprodigiosin) in

XX

CC Streptomyces species, particularly Streptomyces coelicolor or
CC Streptomyces lividans. The present sequence is that of a region of the
CC Streptomyces coelicolor genome, which encodes the ScbA, ScbB and ScbR
CC proteins and which is related to the invention.
XX
SQ Sequence 4346 BP; 650 A; 1561 C; 1519 G; 616 T; 0 U; 0 Other;

Query Match 32.8%; Score 229; DB 10; Length 4346;
Best Local Similarity 60.8%; Pred. No. 2.1e-30;
Matches 373; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

Qy 34 CAGGAACGGGCGCTCGCACGCGGAGGCGATCGTGGGCGAGCCGCTCGGTCTTCGAC 93
Db 2270 CAGGACCGGGCGATCCGACGCGGAGAGATCTCTGGAGCGCGCGGAGGTTCTTCGAG 2329

Qy 94 GAGTACGGTTTCGAGCGCCACAGTGGCAGAGATCCTCTCGGGGCGCTCGGTTCACCAAG 153
Db 2330 AAGCAGGGCTACCAAGTGCACGATCACGGAGATCTCAAGTGGCGCGGTGACCAAG 2389

Qy 154 GGGCGGATGATTCCTTCCATTCGTTCCAAAGGAAGAGCTGGCGCGCGGCTGCTGGCGGAG 213
Db 2390 GGAGCGCTTCTTCCATTCGTTCCAAAGGAAGAGCTGGCGCTGGCGGCTTTCGAGCC 2449

Qy 214 CAGACCTCGACGTGGCGGTGCGGAATCCGCTCCAAAGGCGCAGGAACCTGTAGACCTC 273
Db 2450 CAGGAACCAACAGGCGCTTCGGAGCAACCCCTCCGGTGCAGAACTCATCGACATG 2509

Qy 274 ACCATGCTGGTCCCGCACCGCATGTGTCACGATCCGATCCTGCGGGCGGCGACCGCGCTC 333
Db 2510 GGCATGTTGTTCTGTACCGCTTGGCACGACGTGTCGCGCGCGCGCGGTGCGCTC 2569

Qy 334 GCATGGACCGGGGCGGTGACATTCCTCGAGCGCAACCCGTTCCGGGAGTGGGGCGAC 393
Db 2570 TCATGGACCGAGCGCGCACGCTCTCGATCGCGAGGACCCCTTCGCTCGCGACGAG 2629

Qy 394 ATCTCGCGCCAGCTCTGCGGAGGACACAGGAGACGGGGGAGGTGCTTCGCACTGTAAC 453
Db 2630 ACATCTCTGAAGTCTGTAACCGAGGCAAGGAGTGTGTCGCCCCATGTGTC 2689

Qy 454 CCGAAAAAGACCGGCGACTTCATCGTGGCTGTGTTCCACCGGGTCCAGGCGGTCTCCCGG 513
Db 2690 ACCACCGACTCGGCGCATCTCTACGTGGGCGACGTTTCGCGGGATACAGGTGCTGCCAG 2749

Qy 514 GTCACCTCCGACCGCGGAGACCTCGGCGCACCGGATCTCGGTGATGTGGAACACGCTGCTG 573
Db 2750 ACGGTACGCGACTACCGAGACCTCGAAACCGCTACGCGCTGCTGCAAGACACATCTG 2809

Qy 574 CCCAGCATGTCGCGCGCTCATGCTGACCTGGAACCGGCGGAGGCGGATCGGG 633
Db 2810 CCGGCATCGCGGTTCCCTCCGCTGTCGCGCGCTGATCTCTCCGAGGAGCGCGAGCA 2869

Qy 634 AAGGTCCGCGCG 646
Db 2870 CGCCTCGCGCG 2882

RESULT 4

AAZ99101/c

ID AAZ99101 standard; DNA; 12905 BP.

XX

XX AAZ99101;

XX

XX 21-JUN-2000 (first entry)

XX

XX S. fradiae tylosin biosynthetic pathway gene cluster antisense strand.

XX

XX Gene cluster; macrolide antibiotic; tylosin; biosynthetic pathway;

XX recombinant; Streptomyces; hybrid antibiotic; ss.

XX

XX Streptomyces fradiae.

XX

XX Key Location/Qualifiers

XX complement (294..968)

XX

XX

FT /tag= a
 FT /label= ORF11
 FT /gene= "tylO"
 FT /product= "regulatory protein"
 FT 1315..3291
 FT /tag= b
 FT /label= ORF10
 FT /gene= "tylP"
 FT /product= "acyl-CoA dehydrogenase"
 FT complement(3452..4069)
 FT /tag= c
 FT /label= ORF9
 FT /gene= "tylO"
 FT /product= "delta-TDP-glucose epimerase"
 FT complement(4106..4873)
 FT /tag= d
 FT /label= ORF8
 FT /gene= "tylP"
 FT /product= "macrosin-O-methyltransferase"
 FT 5137..6399
 FT /tag= e
 FT /label= ORF7
 FT /gene= "tylH1"
 FT /product= "cytochrome P450"
 FT 6435..6680
 FT /tag= f
 FT /label= ORF6
 FT /gene= "tylH2"
 FT /product= "ferrodoxin"
 FT /note= "the termination codon of this ORF overlaps with
 FT the initiation codon of ORF5"
 FT 6677..7687
 FT /tag= g
 FT /label= ORF5
 FT /gene= "tylD"
 FT /product= "delta-TDP-glucose 4-6 dehydratase"
 FT /note= "the initiation codon of this ORF overlaps with
 FT the termination codon of ORF6"
 FT 7705..8892
 FT /tag= h
 FT /label= ORF4
 FT /gene= "tylE"
 FT /product= "methyltransferase"
 FT 8938..10206
 FT /tag= i
 FT /label= ORF3
 FT /gene= "tylN"
 FT /product= "glycosyl transferase"
 FT complement(10276..11118)
 FT /tag= j
 FT /label= ORF2
 FT /gene= "tlrB"
 FT /product= "tylosin resistance protein"
 FT 11475..12671
 FT /tag= k
 FT /label= ORF1
 FT /gene= "ddcA"
 FT /product= "D-alanine carboxypeptidase"

WO200009168-A1.

17-FEB-2000.

02-AUG-1999; 99WO-ES000248.

05-AUG-1998; 98ES-00001682.

(ANTI) ANTIBIOTICOS SAU.

XX Fouces Martinez R, Mellado Duran E, Diez Garcia B;
 PI Esteban Morales M, Bernasconi E, Barredo Fuente JL;
 XX WPI; 2000-205718/18.

DR

DR P-PSDB; AAY83784, AAY83785, AAY83786, AAY83787, AAY83788, AAY83789,
 DR AAY83790, AAY83791, AAY83792, AAY83793, AAY83794.
 XX
 FT Increasing tylosin and/or hybrid antibiotic production in Streptomyces
 FT species by introducing genes of the Streptomyces fradiae tylosin
 FT biosynthetic cluster.
 XX
 PS Claim 7; Page 33-43; 63pp; Spanish.
 XX
 CC This sequence corresponds to the antisense strand for the gene cluster
 CC encoding proteins involved in the macrolide antibiotic tylosin
 CC biosynthetic pathway. The invention relates to methods of increasing the
 CC level of production of tylosin and/or hybrid antibiotics in Streptomyces
 CC spp. comprising culturing cells that have been transformed with a vector
 CC containing a Streptomyces fradiae genomic DNA fragment. The method is
 CC useful either for producing recombinant S. fradiae strains that
 CC overproduce the antibiotic tylosin or for producing recombinant strains
 CC of other Streptomyces spp. that will produce novel hybrid antibiotics
 XX
 SQ Sequence 12905 BP; 1859 A; 4723 C; 4547 G; 1776 T; 0 U; 0 Other;

Query Match 18.0%; Score 126; DB 3; Length 12905;
 Best Local Similarity 56.2%; Pred. No. 7.4e-13;
 Matches 342; Conservative 0; Mismatches 250; Indels 16; Gaps 5;
 QY 28 GTGCGACAGGAACGGCGGTCGCGACGGGAGGCGATCGTGGGCGACCGCCTCGGTC 87
 DB 965 GCGCGACAGGAACGGCGGCGCCAGACCGCGCGAGCATGT---CGCGCGCGCGCGG 909
 QY 88 TTCGACGAGTACGGTTGAGGCCGCCACAGTGGCAGAGATCCTCTCGCGGGCCTCGGTC 147
 DB 908 TTCGACGAACTGGGCTAGAGCGGACCACTATCGGGAGATCTCAAGAGGTCCGGGTG 849
 QY 148 ACCAAGGGCGGATGTACTTCCACTTCCAAAGGAAGAGCTGSCCGCGCGCGTCG 207
 DB 848 ACCAAGGGAGCGCTGTACTTCCACTTACGTCCAGGAGCAGCTCGCGAGGAAGTGTG 789
 QY 208 GCGGACGAGACCCCTGCACGTG---GCGGTGCGCGAATCCGGCTCCAAGCGCAGGAAC 264
 DB 788 ACGAGTCAGCTCCGGGCGGTGCCCCGGTGGAGGAACAGCAGCTGCTCTCCAGCAGATC 729
 QY 265 GTAGACCTCACCATGCTGTGTCGCCA---CGGCATGCTGCAGATCCGATCTCTCGGGCG 321
 DB 728 ATCGACGAGACGCTGCTGTGTCGCCAAGTGTCTCAGCAGGGCGATCCGCTGGTGGCGGC 669
 QY 322 GGCACGCGGCTCGCACTGGACCCAGGGGC-----GGTGGACTTCTCCGACGCCAACCCG 375
 DB 668 AGTGTCCGCTCAGGTGAGCGCGCGGCCCGCCAGGAGCGGCTGGACCGCAGGGCGCG 609
 QY 376 TTCGGCGAGTGGGGGACATCTGCGCCAGCTCTCTGGCGGAGGACAGGAACGGGGGAG 435
 DB 608 ATGCAGGAGTGGATCGGCCACGGCAGGATCTCTCAGAAGGGCGAGCGCGGTGAA 549
 QY 436 GTGCTTCGCGACGTGAACCCGAAAGACCGCGGACTTCATGTCGGGCTGCTTACCGGG 495
 DB 548 CTGCTGCGCGGCTCGATGTGCACGCGGTGGCGAGAAATGCTGTCGGCGGTTTACCCGT 489
 QY 496 CTCAGGGGCTCTCCGGGTCACTCCGACCGCAGGACCTCGGCCACCGGATCTCGGTG 555
 DB 488 GCCCAGATCTGTGAAACATCTGACCGGGCAGCGCGATCTGCTGGAGCGGTGACCGAC 429
 QY 556 ATGTGGAACACGCTGCTGCCAGCATCGTGC CGGGCTCCATCTGACTGAGCTCGAAACC 615
 DB 428 ATGCACCGGCATCTCATGACCTCGGTC-CGGTGCCGGCGGTGCTGCTGGTGGCGGCTTC 370
 QY 616 GCGGAGGA 623
 DB 369 TCCGCGGA 362

RESULT 5

AAZ99100

ID AAZ99100 standard; DNA; 12905 BP.

```

XX AC AAZ99100;
XX DT 21-JUN-2000 (first entry)
XX DE S. fradiae tylosin biosynthetic pathway gene cluster sense strand.
XX DT
XX DE Gene cluster; macrolide antibiotic; tylosin; biosynthetic pathway;
XX DT recombinant; Streptomyces; hybrid antibiotic; ss.
XX DE Streptomyces fradiae.
XX DT
XX DE Location/Qualifiers
XX DT complement(235..1431)
XX DT /tag= a
XX DT /label= ORF1
XX DT /gene= "ddca"
XX DT /product= "D-alanine carboxypeptidase"
XX DT 1788..2630
XX DT /tag= b
XX DT /label= ORF2
XX DT /gene= "tlrB"
XX DT /product= "tylosin resistance protein"
XX DT complement(2700..3968)
XX DT /tag= c
XX DT /label= ORF3
XX DT /gene= "tylN"
XX DT /product= "glycosyl transferase"
XX DT complement(4014..5201)
XX DT /tag= d
XX DT /label= ORF4
XX DT /gene= "cyle"
XX DT /product= "methyltransferase"
XX DT complement(5219..6229)
XX DT /tag= e
XX DT /label= ORF5
XX DT /gene= "tylD"
XX DT /product= "delta-TDP-glucose 4-6 dehydratase"
XX DT /notes= "the initiation codon of this ORF overlaps with
XX DT the termination codon of ORF6"
XX DT complement(6226..6471)
XX DT /tag= f
XX DT /label= ORF6
XX DT /gene= "tylH2"
XX DT /product= "ferrodoxin"
XX DT /notes= "the termination codon of this ORF overlaps with
XX DT the initiation codon of ORF5"
XX DT 6507..7769
XX DT /tag= g
XX DT /label= ORF7
XX DT /gene= "tylH1"
XX DT /product= "cytochrome P450"
XX DT 8033..8800
XX DT /tag= h
XX DT /label= ORF8
XX DT /gene= "cylF"
XX DT /product= "macrosin-O-methyltransferase"
XX DT 8837..9454
XX DT /tag= i
XX DT /label= ORF9
XX DT /gene= "tylO"
XX DT /product= "delta-TDP-glucose epimerase"
XX DT complement(9615..11591)
XX DT /tag= j
XX DT /label= ORF10
XX DT /gene= "tylP"
XX DT /product= "acyl-CoA dehydrogenase"
XX DT 11938..12612
XX DT /tag= k
XX DT /label= ORF11
XX DT /gene= "tylQ"
XX DT /product= "regulatory protein"

```

```

PN WO200008168-A1.
XX
XX 17-FEB-2000.
XX
XX 02-AUG-1999; 99WO-ES000248.
XX
XX 05-AUG-1998; 98ES-00001682.
XX
XX (ANTI ) ANTIBIOTICOS SAU.
XX
XX Fouces Martinez R, Mellado Duran E, Diez Garcia B;
XX Esteban Morales M, Bernasconi E, Barredo Fuente JL;
XX
XX WPI; 2000-205718/18.
XX
XX P-PsDB; AAY83784, AAY83785, AAY83786, AAY83787, AAY83788, AAY83789,
XX AAY83790, AAY83791, AAY83792, AAY83793, AAY83794.
XX
XX Increasing tylosin and/or hybrid antibiotic production in Streptomyces
XX species by introducing genes of the Streptomyces fradiae tylosin
XX biosynthetic cluster.
XX
XX Claim 7; Page 26-32; 63pp; Spanish.
XX
XX This sequence corresponds to the sense strand of the gene cluster
XX encoding proteins involved in the macrolide antibiotic tylosin
XX biosynthetic pathway. The invention relates to methods of increasing the
XX level of production of tylosin and/or hybrid antibiotics in Streptomyces
XX spp. comprising culturing cells that have been transformed with a vector
XX containing a Streptomyces fradiae genomic DNA fragment. The method is
XX useful either for producing recombinant S. fradiae strains that
XX overproduce the antibiotic tylosin or for producing recombinant strains
XX of other Streptomyces spp. that will produce novel hybrid antibiotics
XX
XX Sequence 12905 BP; 1776 A; 4548 C; 4722 G; 1859 T; 0 U; 0 Other;
XX
XX Query Match 18.0%; Score 126; DB 3; Length 12905;
XX Best Local Similarity 56.2%; Pred. No. 7.4e-13;
XX Matches 342; Conservative 0; Mismatches 250; Indels 16; Gaps 5;
XX
XX QY 28 GTGCGACAGAAACGGGCCGTCGCGACGGCGAGGCGTGGTGGGCGAGCGCGCTCGGTC 87
XX DB 11941 GCGCGACAGAAACGGGCCGTCGCGACGGCGAGGCGTGGTGGGCGAGCGCGCTCGGTC 11997
XX
XX QY 88 TTGCGACGAGTACGGGTTGAGGCGCGCCACAGTGGGAGAGATCTCTTCGGGGGCTCGGTC 147
XX DB 11998 TTGCGACGAACTGGGTTACGAGGCGCACCATCGCGAGATCTCAAGAGTCCGGGGTG 12057
XX
XX QY 148 ACCAAGGGCGGATGTACTTCCACTTCGCTTCCAGGAGAGCTGGCCCGCGCGCTGCTG 207
XX DB 12058 ACCAAGGGGAGCGTGTACTTCCACTTCCAGTCCAGGAGAGCTCGCGGAGGAGTGCTG 12117
XX
XX QY 208 GCGGAGCAGACCTTGACCGTG---GCGGTGCGCGGAATCCGGGTCCAGGCGCAGGAATCG 264
XX DB 12118 ACGAGTCAGCTCGGGCGCGTCCCGCGGTGGAGAACAGCGACTCGCTCTCCAGCAGATC 12177
XX
XX QY 265 GTAGACCTCACCATGCTGTGTCGCCA---CGGATGCTGCAGATCCGATCTTCGCGGGCG 321
XX DB 12178 ATCGACGAGACGCTGCTGTGCGCCAACTGCTCAGCAAGGGCGATCCGCTGTGTCGCGG 12237
XX
XX QY 322 GGCACGCGGCTCGACTGGACACAGGGGC-----GGTGGACTTCTCCGACGCCACCCCG 375
XX DB 12238 AGTGTCCGGCTCACGTTGGAGCGCGCGCGCCCGCCAGGAGCGGCTTGACCGAGGGCGCG 12297
XX
XX QY 376 TTGCGCGAGTGGGGGACATCTGCGCCCGAGCTCCTGGCGGAGGACAGCAAGACGGGGGAG 435
XX DB 12298 ATGCAAGAGTGGATCGGCACCGCAGGAGTCTCTCAGAAAGCGCGGCGGCGTGA 12357
XX
XX QY 436 GTGCTTCGCACTGTAACCCGAAAGACCGCGGACTTCATCGTGGGCTGTTTCACCGGG 495
XX DB 12358 CTGCTGCCGCGCTCGATGTCACGCGCGTGGGAGAAATCTGTGTCGCGGTTTCACCGGT 12417
XX
XX QY 496 CTCAGGCGGTCTCCCGGGTCACTCCGACCGCCAGGACCTCGGCCACCGATCTCGGTG 555

```

Db 12418 GCCAGATCTGTGCAACATCTGACGGGCGACGCCGATCTGCTGGAGCGGTGACCGAC 12477
QY 556 ATGTGGAACACGTCGTCGCCAGCATGCGCGGTCCTCATGCTGACCTGATGCAAAACC 615
Db 12478 ATGCACCGGCATCTCATGACCTCGGTC-CGGTCCCGCGCTGCTGTGGCGGCTGACATTC 12536
QY 616 GCGGAGGA 623
Db 12537 TCCGCGGA 12544

RESULT 6
ABV75877/c
ID ABV75877 standard; DNA; 1819 BP.
XX
AC ABV75877;
XX
DT 07-FEB-2003 (first entry)
XX
DE Streptomycetes coelicolor type II thioesterase scot gene.
XX
KW Thioesterase; enzyme; Scot; polyketide; antibiotic; gene; ds.
XX
OS Streptomycetes coelicolor.
XX
FH Location/Qualifiers
FT 748..1554
FT /*tag= a
FT /product= "Type II thioesterase"
FT /note= "a polynucleotide comprising the CDS is
FT specifically claimed in Claim 10"

WO200279461-A2.

10-OCT-2002.

28-MAR-2002; 2002WO-PL000027.

28-MAR-2001; 2001PL-00346786.

(KUCZ/) KUCZEK K.
(KOTO/) KOTOWSKA M.
(PAWL/) PAWLICK K.
(WIERN/) WIERNIK D.

Kuczek K, Kotowska M, Pawlik K, Wiernik D;

WPI; 2003-058432/05.
P-PSDB; ABP55143.

New thioesterase type II (TE II) protein, useful for expression of a protein having TE II activity or for increasing biosynthetic efficiency of a polyketide synthase complex.

Claim 1; Page 11-12; 13pp; English.

The present sequence is that of the type II thioesterase (TE II) scot gene of Streptomycetes coelicolor A3(2) M145. The gene was isolated from a genomic library using a probe based on the central region of the acyltransferase domain from type I polyketide synthase, including the active site region. DNA reacting with the probe, obtained as 2 restriction fragments, was cloned and sequenced. Correction of frame-shift mistakes provided the present sequence, which includes an open reading frame (also claimed), designated scot. This encodes a 268-amino acid protein (see ABP55143) having a mol.wt. of 28,686 and a pI of 6.17, which has sequence similarity to known TE II enzymes. The isolated gene also includes a region active in transcriptional regulation of scot. The scot enzyme is active in polyketide, especially macrolide, biosynthesis when associated with a multienzyme complex of a polyketide synthase. Its relatively broad substrate tolerance will facilitate the production of novel, heterologous polyketide synthase complexes. The scot gene may be transferred to a different streptomycete to increase the efficiency of production of a polyketide (possibly an antibiotic)

XX SQ Sequence 1819 BP; 246 A; 683 C; 620 G; 270 T; 0 U; 0 Other;
Query Match 12.7%; Score 88.6; DB 8; Length 1819;
Best Local Similarity 54.2%; Pred. No. 2.1e-06;
Matches 215; Conservative 0; Mismatches 164; Indels 18; Gaps 1;
QY 34 CAGGAACCGGCGCTCCGACGCGGCGAGCTGTCGGGGCAGCGCCTCGGTCTTCGAC 93
Db 500 CAGGAGCGGCGGCGAGTACGGCTCGAGCTCTGATTCTCTCGCGCGCGAGGTGTCGAC 441
QY 94 GAGTACGGGTTGAGGCGCGCCACAGTGGCAGAGATCTCTCGCGGGCTCGGTACCAAG 153
Db 440 CAGGAGGGGTTTCGGCGCGCGCTCGCTCACGATGATCAGTCCCGGCGGGGTGAGCAAC 381
QY 154 GCGCGATGTACTTCCACTTCCGCTTCCAAAGGAAGAGCTGGCCCGCGGTGCTGCGCGA- 212
Db 380 GGGCACTGCACTTCCACTTCCGACAAAGAACGCGGTGGCGAGCGGTGAGGGCGAG 321
QY 213 -----GCAGACCTGTCACGCTGGCGGTGCCGGAATCCGGTCCAAAGCG 255
Db 320 GCGTTGTGCTGCTCCGGCAGATGCCACGCTGGCGGAGGGGCGCACGCCCTCGCTC 261
QY 256 CAGGAACCTGTAGACTCACCATGCTGTGCTGCCACGCGATGCTGCAGATCCGATCTCTG 315
Db 260 CAGTCCCTCGTCGACACCTCGCACGCTGCGCCCGACGACTCCAGGACGATGTCGTGTG 201
QY 316 CCGGCGGCGACGCGGCTCGCACTGGACCCAGGGGGCGGTGGAGCTTCTCGACGCCAACCCG 375
Db 200 CCGGCGGCTTTCGGGCTGAGGGGACACGACCTGGAAGGACGCGCGGATGCGCGCG 141
QY 376 TTCGCGAGTGGGCGACATCTGCCCGCCAGTCTCTGG 412
Db 140 CACTGGGTGACTGGGTGAGCAGCGGCTGACCGTGG 104

RESULT 7

AAS07627

ID AAS07627 standard; DNA; 9521 BP.

AC AAS07627;

XX 11-SEP-2003 (revised)

DT 23-OCT-2001 (first entry)

Streptococcus coelicolor genomic DNA from plasmid SCP1.

SCP1; ds; methylenomycin cluster; mmc; MmyR; MmfP; MmfH; MmfL; MmfR; MmyT; MmyO; MmyG; MmyU; Mmr; heterologous gene expression.

Streptomycetes coelicolor.

Key Location/Qualifiers
CDS complement(796..1407)

FT /*tag= a
FT /product= "MmyR"
FT /note= "This sequence is specifically claimed"

FT /partial
FT /note= "The mmyR polypeptide may be encoded from the ATG
FT at 1387-1399"
FT complement(1390..1557)

promoter

FT /*tag= b
FT /label= mmyR_promoter

FT /note= "This sequence is specifically claimed, and
FT optionally excludes nucleotides 1390-1409"
FT complement(1558..2355)
FT /*tag= c

FT /product= "MmfP"
FT /note= "This sequence is specifically claimed"

FT /partial
FT /note= "The mmfP polypeptide may be encoded from the ATG
FT at 2350-2352"
FT complement(2352..3554)
FT CDS


```
FT FT /*tag= d
FT FT /product= "MmFh"
FT FT /note= "This sequence is specifically claimed"
FT FT complement(3551..4612)
FT FT /tag= e
FT FT /product= "MmFL"
FT FT /note= "This sequence is specifically claimed"
FT FT 4613..4806
FT FT /tag= f
FT FT /label= "MmFR_promoter"
FT FT /note= "This sequence is specifically claimed"
FT FT complement(4613..4806)
FT FT /tag= g
FT FT /label= "MmFL/mmFHP_promoter"
FT FT /note= "This sequence is specifically claimed"
FT FT 4807..5451
FT FT /tag= h
FT FT /product= "MmFR"
FT FT /note= "This sequence is specifically claimed"
FT FT 5452..5675
FT FT /tag= i
FT FT /label= "MmYOG_promoter"
FT FT /note= "This sequence is specifically claimed"
FT FT 5676..6401
FT FT /tag= j
FT FT /product= "MmYT"
FT FT /note= "This sequence is specifically claimed"
FT FT 6432..7553
FT FT /tag= k
FT FT /product= "MmYO"
FT FT /note= "This sequence is specifically claimed"
FT FT 7536..8817
FT FT /tag= l
FT FT /product= "MmYG"
FT FT /note= "This sequence is specifically claimed"
FT FT complement(8780..9151)
FT FT /tag= m
FT FT /product= "MmYJ"
FT FT /partial
FT FT /note= "The mmyJ polypeptide may be encoded from the ATG
FT FT at 9113-9115"
FT FT 9333..9521
FT FT /tag= n
FT FT /product= "Mmr"
FT FT /partial
FT FT /note= "No stop codon"
FT FT
XX WO200148228-A1.
XX
XX 05-JUL-2001.
XX
XX 20-DEC-2000; 2000WO-GB004972.
XX
XX 23-DEC-1999; 99GB-00030477.
XX
XX (PLAN-) PLANT BIOSCIENCE LTD.
XX
XX Chater KF, Bruton CJ, O'rouke SJ, Wietzorrek AW;
XX
XX WPI; 2001-425675/45.
XX
XX P-PSDB; AAU04039, AAU04040, AAU04041, AAU04042, AAU04043, AAU04044,
XX AAU04045, AAU04046.
XX
XX Novel expression cassette for expressing a nucleic acid of interest,
XX derived from the regulatory region of methylenomycin gene cluster of SCP1
XX plasmid of Streptomyces coelicolor A3(2).
XX
XX Example 3; Fig 7; 142pp; English.
XX
XX The sequence represents an expression cassette (present on plasmid SCP1)
XX and is the regulatory region of the methylenomycin cluster (mnc) from
XX Streptomyces coelicolor A3(2), which encodes the MmyR, MmFp, MmFL,
XX MmFR, MmYT, MmYO, MmYG, MmYU and partial Mmr polypeptides. The expression
```

```
CC cassette is useful for expressing a nucleic acid of interest,
CC substantially only when the host cell culture reaches high cell density
CC at or close to the stationary phase of host cell culture. In particular
CC the system is useful in regulating methylenomycin production. Reduced or
CC no expression of the nucleic acid of interest is observed earlier in
CC growth, avoiding toxic effects of some gene products on growth and the
CC system does not require addition of exogenous inducer. The methylenomycin
CC cluster naturally present on a highly transmissible plasmid permits
CC properly regulated expression in diverse Streptomyces host and the
CC expression is driven by a strong promoter, leading to high yield of the
CC desired end product. (Updated on 11-SEP-2003 to standardise OS field)
XX
XX SQ Sequence 9521 BP; 1399 A; 3223 C; 3416 G; 1493 T; 0 U; 0 Other;
```

```
Query Match 12.3%; Score 85.8; DB 4; Length 9521;
Best Local Similarity 50.1%; Pred. No. 5.5e-06;
Matches 281; Conservative 0; Mismatches 262; Indels 18; Gaps 2;

QY 30 GCACAGGAACGGGCGGTCCGCACGGCGGAGCGGATCGTGGGGGACGCCCTCGGTCTT 89
DB 4872 GCAGCAGGAGCGGTCTGATCAAGACCGGGCCAGATCTTGGAGGCGCGTTCGAGATCTT 4931
QY 90 CGACGAGTACGGGTTCCGAGCGCCACAGTGGCAGAGATCCTCTCGGGGCGCTCGGTTCAC 149
DB 4932 CGGTCCGCGGCTACCGAGGGGCTCCGTCAAGGACGTTGCCGAGCGGTTCGGCATGAC 4991
QY 150 CAAGGGCGCGATGTACTTCCACTTCCCTTCCAAAGGAAGAGCTGGCCCGGGCGGTGCTGGC 209
DB 4992 CAAGGGCGGGTGTACTTCCACTTCCCGCAGCAAGGAATCCTGGCCATCGCCGTGGTGA 5051
QY 210 CGAGCAGACCTCGACGTGGCGG-----TGCCGAATCCGGCTCCAAGGC----- 254
DB 5052 GGAGCACTACGGCGCGTGGCCCGCAGCATGGAAGAGATCCGCATCCAGGGCTTCACACC 5111
QY 255 ---GCAGGAACCTGGTAGACCTCACCATGCTGGTCCGCCACGCGATCTGCACGATCCGAT 311
DB 5112 GCTGGAGACGCTCGAGGAGATGCTCATCGCGGGCGCAGGCTTCGCGACACGCCCGT 5171
QY 312 CCTCGGGGGGCGACGCGCTCGCATGGAACGAGGGGGGGGTGGACTTCTCCGACGCCAA 371
DB 5172 GATGAGCGCGGTGCCCGGCTGCAGAGTGCAGCGCGCTTCATCGACGGAGCTGCCCT 5231
QY 372 CCGCTTCGCGAGTGGGGGACATCTGCCCGCAGCTCTTGGCGGAGGACAGGAACGGGG 431
DB 5232 GCCCTACGTGGACTGGACCCACCTGCTGGAGGTGCGGTTGCAGGACGCCCGTGAGGCCG 5291
QY 432 GGAGGTGCTTCGCGACGTGAACCCGAAAGACCGGCGACTTCATCGTGGCTGCTTCAC 491
DB 5292 CCAGTTGGGGGGGCGTGCATCCCGCAGCAGCTGCCCGCTTCCTGGTGGCGCCCTTCTT 5351
QY 492 CGGGCTCCAGGGCGGTCTCCCGGGTCACTCCGACCGCCAGGACCTCGGCCACCGGATCTC 551
DB 5352 CGGCATGCAGCAGCTCTCCGACAAATCGCACCGAGCGGACATCATGAGCGGTGGCA 5411
QY 552 GGTGATGTGAACACCGTCT 572
DB 5412 GGAGCTGGCGGAGCTGATGTT 5432
```

```
RESULT 8
AAV68520
ID AAV68520 standard; DNA; 4257 BP.
XX
XX AAV68520;
AC
XX 29-JAN-1999 (first entry)
DT
XX
XX DE The nucleotide sequence of the Herpes simplex virus ICP4.
XX
XX KW Herpes simplex virus ICP4; HSV ICP4; modulator; apoptosis; stimulation;
XX inhibition; HSV infection; ss.
XX
XX OS Herpes simplex virus unknown type.
```

```
XX Key Location/Qualifiers
FH CDS
FT 361..4257
FT /*tag= a
FT /product= "HSV ICP4 protein"
XX WO9846637-A2.
XX
XX 22-OCT-1998.
XX
XX 16-APR-1998; 98WO-US007573.
XX
XX 16-APR-1997; 97US-00843659.
XX
XX (ARCH-) ARCH DEV CORP.
XX
XX Leopardi R, Roizman B;
XX
XX WPI; 1998-594559/50.
XX
XX P-PSDB; AAW80810.
XX
XX Use of herpes simplex virus U(s)3 polypeptide - for developing products
XX for modulating apoptosis in cells and for identifying compounds which act
XX as stimulators or inhibitors of apoptosis.
XX
XX Example 2; Page 60-63; 85pp; English.
XX
XX This is the nucleotide sequence of Herpes simplex virus ICP4 used in the
XX method of the invention as modulators of apoptosis. The methods and
XX products can be used to identify compounds which modulate (stimulate or
XX inhibit) apoptosis in cells. They can be used to immortalise cells for
XX the study of these cells or for growing cells in large numbers for the
XX productions of proteins. They can also be used for stimulating apoptosis
XX in cells, e.g. for treating a subject with a HSV infection
XX
XX Sequence 4257 BP; 412 A; 1768 C; 1663 G; 414 T; 0 U; 0 Other;
XX
XX Query Match 11.2%; Score 78.2; DB 2; Length 4257;
XX Best Local Similarity 45.0%; Pred. No. 0.00012;
XX Matches 293; Conservative 0; Mismatches 358; Indels 0; Gaps 0;
XX
XX 43 GCCGTCCGACGCGGAGCGATCGTGGGAGCGCGCTCGTTCGACGAGTACGGG 102
XX 1867 GCGGTGCGCGCGCGCGACCGCGGAGCGCGGTGCGCGGCGCGCGCGCGCGG 1926
XX
XX 103 TTGAGGCGCGCACAGTGCAGATCTCTCGCGGCGCTCGTCAACGAGCGCGATG 162
XX 1927 ATCTCGCGCGCGTGGGCGGTGTCTCGCGCGCGCGCTCCCGCGCGCGCGCGG 1986
XX
XX 163 TACTTCCACTTCCAGGAGAGCTGGCGCGCGGTGTGGCGCGGAGCGACCTGTG 222
XX 1987 CCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 2046
XX
XX 223 CAGTGGCGCGCGCGGAAATCCGCGCTCAAGGCGCGAGAACTGTGTAGACCTACATGCTG 282
XX 2047 CGCGTGGCGCGTGCAGTGCCTGCGCGCGCTGCGCGCGGATCTGTGAGCGCTGGCGGAGGCG 2106
XX
XX 283 GTGCGCGCGGATGTCGACGATCCGATCTGCGCGCGCGCGCGCGCGCGCGCGCG 342
XX 2107 TTGAGCGCGCGCTGCGCGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2166
XX
XX 343 CAGGCGCGCGGTGACCTTCTCCGACGCGAAACCGCTTTCGCGGAGTGGGCGGACATCTCGCGC 402
XX 2167 CGGCGCGGAGGACCGCGCGCGCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGG 2226
XX
XX 403 CAGCTCTGCGGAGGACAGGAAACGCGGCGGAGTGTCTTCCGCGCGTGTACCGGAAAG 462
XX 2227 CTGCGCGCGTGTGCGCGCGTGTGCGCGCGCGTGTGCGCGCGCGCGCGCGCGCGCG 2286
XX
XX 463 ACCGCGGACCTTCACTGTCGCGCTCTTCAACCGCGCTCCAGCGCGTCTCCGCGGTCACCTCG 522
XX 2287 CGCGGCGACCTGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2346
XX
XX Location/Qualifiers
XX CDS
XX 361..4257
XX /*tag= a
XX /product= "HSV ICP4 protein"
XX WO9846637-A2.
XX
XX 22-OCT-1998.
XX
XX 16-APR-1998; 98WO-US007573.
XX
XX 16-APR-1997; 97US-00843659.
XX
XX (ARCH-) ARCH DEV CORP.
XX
XX Leopardi R, Roizman B;
XX
XX WPI; 1998-594559/50.
XX
XX P-PSDB; AAW80810.
XX
XX Use of herpes simplex virus U(s)3 polypeptide - for developing products
XX for modulating apoptosis in cells and for identifying compounds which act
XX as stimulators or inhibitors of apoptosis.
XX
XX Example 2; Page 60-63; 85pp; English.
XX
XX This is the nucleotide sequence of Herpes simplex virus ICP4 used in the
XX method of the invention as modulators of apoptosis. The methods and
XX products can be used to identify compounds which modulate (stimulate or
XX inhibit) apoptosis in cells. They can be used to immortalise cells for
XX the study of these cells or for growing cells in large numbers for the
XX productions of proteins. They can also be used for stimulating apoptosis
XX in cells, e.g. for treating a subject with a HSV infection
XX
XX Sequence 4257 BP; 412 A; 1768 C; 1663 G; 414 T; 0 U; 0 Other;
XX
XX Query Match 11.2%; Score 78.2; DB 2; Length 4257;
XX Best Local Similarity 45.0%; Pred. No. 0.00012;
XX Matches 293; Conservative 0; Mismatches 358; Indels 0; Gaps 0;
XX
XX 43 GCCGTCCGACGCGGAGCGATCGTGGGAGCGCGCTCGTTCGACGAGTACGGG 102
XX 1867 GCGGTGCGCGCGCGCGACCGCGGAGCGCGGTGCGCGGCGCGCGCGCGCGG 1926
XX
XX 103 TTGAGGCGCGCACAGTGCAGATCTCTCGCGGCGCTCGTCAACGAGCGCGATG 162
XX 1927 ATCTCGCGCGCGTGGGCGGTGTCTCGCGCGCGCGCTCCCGCGCGCGCGG 1986
XX
XX 163 TACTTCCACTTCCAGGAGAGCTGGCGCGCGGTGTGGCGCGGAGCGACCTGTG 222
XX 1987 CCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 2046
XX
XX 223 CAGTGGCGCGCGCGGAAATCCGCGCTCAAGGCGCGAGAACTGTGTAGACCTACATGCTG 282
XX 2047 CGCGTGGCGCGTGCAGTGCCTGCGCGCGCTGCGCGCGGATCTGTGAGCGCTGGCGGAGGCG 2106
XX
XX 283 GTGCGCGCGGATGTCGACGATCCGATCTGCGCGCGCGCGCGCGCGCGCGCGCG 342
XX 2107 TTGAGCGCGCGCTGCGCGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2166
XX
XX 343 CAGGCGCGCGGTGACCTTCTCCGACGCGAAACCGCTTTCGCGGAGTGGGCGGACATCTCGCGC 402
XX 2167 CGGCGCGGAGGACCGCGCGCGCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGG 2226
XX
XX 403 CAGCTCTGCGGAGGACAGGAAACGCGGCGGAGTGTCTTCCGCGCGTGTACCGGAAAG 462
XX 2227 CTGCGCGCGTGTGCGCGCGTGTGCGCGCGCGTGTGCGCGCGCGCGCGCGCGCGCG 2286
XX
XX 463 ACCGCGGACCTTCACTGTCGCGCTCTTCAACCGCGCTCCAGCGGTCTCCGCGGTCACCTCG 522
XX 2287 CGCGGCGACCTGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2346
XX
XX ID AAV10362 standard; cDNA; 4257 BP.
XX
XX AC AAV10362;
XX
XX 30-JUN-1998 (first entry)
XX
XX DE Infected cell protein number 4 alpha-4 gene.
XX
XX KW Infected cell protein number 4; ICP4; alpha-4; cell apoptosis;
XX therapeutics; ss.
XX
XX OS Herpes simplex virus unknown type.
XX
XX FH Key Location/Qualifiers
XX FT CDS 361..4257
XX FT /*tag= a
XX FT /product= "Infected cell protein"
XX
XX PN WO9804709-A2.
XX
XX PD 05-FEB-1998.
XX
XX PF 23-JUL-1997; 97WO-US012904.
XX
XX PR 26-JUL-1996; 96US-00690473.
XX
XX PA (ARCH-) ARCH DEV CORP.
XX
XX PI Leopardi R, Roizman B;
XX
XX WPI; 1998-130697/12.
XX
XX P-PSDB; AAW40200.
XX
XX Use of herpes simplex virus ICP4 polypeptide - useful for, e.g. blocking
XX apoptosis in cells, production of proteins and gene therapy.
XX
XX Disclosure; Fig 2; 63pp; English.
XX
XX The herpes simplex virus (HSV) alpha-4 gene encodes infected cell protein
XX number 4 (ICP4). Both the alpha-4 gene and ICP4 can be used to block cell
XX apoptosis. Similarly the administration of an agent that inhibits ICP4 or
XX the alpha-4 gene can induce apoptosis in HSV infected cells. This can be
XX used for the immortalisation of cells, production of proteins, gene
XX therapy, or inhibition of cell death induced in vivo. They can also be
XX used for production of therapeutics comprising inhibitors of HSV ICP4
XX function, useful for treating HSV function
XX
XX Sequence 4257 BP; 412 A; 1768 C; 1663 G; 414 T; 0 U; 0 Other;
XX
XX Query Match 11.2%; Score 78.2; DB 2; Length 4257;
XX Best Local Similarity 45.0%; Pred. No. 0.00012;
XX Matches 293; Conservative 0; Mismatches 358; Indels 0; Gaps 0;
XX
XX 43 GCCGTCCGACGCGGAGCGATCGTGGGAGCGCGCTCGTTCGACGAGTACGGG 102
XX 1867 GCGGTGCGCGCGCGCGACCGCGGAGCGCGGTGCGCGGCGCGCGCGCGCGG 1926
XX
XX 103 TTGAGGCGCGCACAGTGCAGATCTCTCGCGGCGCTCGTCAACGAGCGCGATG 162
XX 1927 ATCTCGCGCGCGTGGGCGGTGTCTCGCGCGCGCGCTCCCGCGCGCGCGG 1986
XX
XX 163 TACTTCCACTTCCAGGAGAGCTGGCGCGCGGTGTGGCGCGGAGCGACCTGTG 222
XX 1987 CCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 2046
XX
XX 223 CAGTGGCGCGCGCGGAAATCCGCGCTCAAGGCGCGAGAACTGTGTAGACCTACATGCTG 282
XX 2047 CGCGTGGCGCGTGCAGTGCCTGCGCGCGCTGCGCGCGGATCTGTGAGCGCTGGCGGAGGCG 2106
XX
XX 283 GTGCGCGCGGATGTCGACGATCCGATCTGCGCGCGCGCGCGCGCGCGCGCGCG 342
XX 2107 TTGAGCGCGCGCTGCGCGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2166
XX
XX 343 CAGGCGCGCGGTGACCTTCTCCGACGCGAAACCGCTTTCGCGGAGTGGGCGGACATCTCGCGC 402
XX 2167 CGGCGCGGAGGACCGCGCGCGCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGG 2226
XX
XX 403 CAGCTCTGCGGAGGACAGGAAACGCGGCGGAGTGTCTTCCGCGCGTGTACCGGAAAG 462
XX 2227 CTGCGCGCGTGTGCGCGCGTGTGCGCGCGCGTGTGCGCGCGCGCGCGCGCGCGCG 2286
XX
XX 463 ACCGCGGACCTTCACTGTCGCGCTCTTCAACCGCGCTCCAGCGGTCTCCGCGGTCACCTCG 522
XX 2287 CGCGGCGACCTGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2346
XX
```

Db 1927 ATCTCGCCGCGCTGGGGCGGTGTCTCCGCGCGCCCGCTCCCGCGGGGGCGACGAC 1986
QY 163 TACTTCCACTTCGCTTCCAAAGAGAGCTGGCCCGCGGGTGTCTGGCCGAGCAGACCCCTG 222
Db 1987 CCGGAGCGCCCGCCACCGCAGCGCGAGCAGCGCGGGCGCGCGCCAGCGCGCG 2046
QY 223 CACGTGGCGGTGCGGAAATCCGGCTCCAAAGGCGCAGAACTGTGTAGACTTCACTATGCTG 282
Db 2047 CGGTGGCGGTGCGAGTGCCTGGCGCGCTGCGCGGATCTGTGAGGCGTGTGGCCGAGGCG 2106
QY 283 GTCCGACGCGATGCTGCAAGATCGATCTGCGGGCGGGCAGCGCGGTCTGCATGAC 342
Db 2107 TTCAGCGGACCTGGCGCGCGTCCCGGGCTGGCGGGCGCGCGCGCGCGCGCGCG 2166
QY 343 CAGGGGCGGTGACTTCTCCACGCAACCCCGTTTCGGCGAGTGGGGCGACATCTCGCGC 402
Db 2167 CGGCGGAGGACCCCGGGCGCGCTTCCTCCGCGCGCGCGCGCGCGCGCGCGCG 2226
QY 403 CAGCTCTGGCGAGGACAGGAAACGGGGGAGGTCTTCCGCACTGTGAACCCGAAAG 462
Db 2227 CTGCGCGGTGGTGGCGAGTGCCTGCTGCGCGCGCGCTGCTGTGCTCATGCGCGCTG 2286
QY 463 ACGGCGACTTATCTGCGCTCTTACGCGGCTCCAGCGGCTCCAGCGGTCTCCGGGTCACTCC 522
Db 2287 CGCGGGGACCTGCGCTGGCGCGCGCGCGAGCGCGCGTGGCGCGCGCGCGCGCG 2346
QY 523 GACCGCGAGACCTCGGCGCACCGGATCTCGGTGATGTGAACACACCTGTGCGCCAGCATC 582
Db 2347 AGCCTGTGCGCGGGCGCTTGGCGCGCGCTGCGCGGACCGCGCGCTGCGAGCTCC 2406
QY 583 GTGCGCGGTCTATGCTGACCTGATGATGAAACCGGCGAGAGCGGATCGGGAAGTTCGG 642
Db 2407 GCGCGCGCGCGCGCGCGACCTGCTGTTGAAACAGAGCGCTGCGCGCGCGCTGCTGGCG 2466
QY 643 GCGGCGCGCGCGCGCGAGTGGCGGCTCGGAGGCTTCGAGGCGCGCTCCGAC 693
Db 2467 GCGGCGCGCGCGCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2517

RESULT 10
ACA37971 standard; DNA; 1224 BP.
XX ID ACA37971 standard; DNA; 1224 BP.
XX AC ACA37971;
XX CT ACA37971;
XX DT 19-JUN-2003 (first entry)
XX DE Prokaryotic essential gene #19628.
XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;
XX KW drug design; gene.
XX KW Mycobacterium avium.
XX OS W0200277183-A2.
XX PN 03-OCT-2002.
XX PD 21-MAR-2002; 2002W0-US009107.
XX PF 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362899P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GI, Yamamoto R, Foreyth RA, Xu HH;
XX WI; 2003-029996/02.

DR P-PSDB; ABU34101.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 14; SEQ ID NO 25841; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1224 BP; 138 A; 451 C; 485 G; 150 T; 0 U; 0 Other;

Query Match 9.8%; Score 68.2; DB 8; Length 1224;
Best Local Similarity 46.4%; Pred. No. 0.0065;
Matches 306; Conservative 0; Mismatches 338; Indels 15; Gaps 2;
QY 36 GGAACGGGCGCTCGCACGCGCGAGTCTGCGCGGCGAGCGCTCTTCGACGA 95
Db 522 GCAGCGGTGTGTCTCGCGGTGCGCGATTTGCGGCTTCCCGCGCGCGAGGTGTGCGGT 581
QY 96 GTACGGGTTCGAGGCGCGCACAGTGGCAGAGATCTCTTCGCGGGCTTCGAGTCAACGAGG 155
Db 582 CCGGTGCGCGAGGCGCATCACCTCGTGGACCGCGCGCGGTGCGCGCGCGCGTGC 641
QY 156 CGCGATGTACTTCCACTTCGCTTCCAAAGAGAGTGGCGCGCGCGGTGCGCGCGAGCA 215
Db 642 GCACCTTCGATTCACCGACGCGCGCGGTGTGTGATGTGTTGCGGCGGTTCGAGGCGCG 701
QY 216 GACCTTGACGTGGCGGTGCGGAAATCCGCTCCAAAGGCGCGAGAACTTGTAGACCTTCA 275
Db 702 GGCCTGCTGAACCGCGCGGTGTCTCGCGCGCGCGCGAGCTGCGCGCGCGCGGTGTC 761
QY 276 CATGCTGTGTGCGCCACGCGCATGCTGCACGATCCGATCTCTGCGCGCGGCGACGCGGTTCG 335
Db 762 CTGTGCTGACGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 821
QY 336 ACTGGAACGAGGGGCGGTGAGCTTCTCCGAC-----GCCAACCGTTCGCGCA 383
Db 822 CCGCGCTTACGTGGCGGTGCTTCTTACCTGGAACCGAATGGACCTTGGCGCGCGCGCGCA 881
QY 384 GTGGCGCGACATCTGCGCGCGCGCGCTCTCTGCGCGAGGCGACAGGAGGAGGTGCTTCC 443
Db 882 CCTGCTGATCTGCGCGGTGCGGGCGGATGACGCTGCGCGCGAGGTGCTGCGCGGTGCC 941

```

QY 444 GCACGTGAACCGAAGACCGCGGACCTTCATCGTCGGTGTCTTCCACCGGCTCCAGGC 503
Db 942 GGCACTACGTACGTCGCGTCCGATCGGCAACGGGAAACAGCGGCTCAACGCGTCCGGT 1001
QY 504 GGTCTCCCGGGTCACTCCGACCGCCAGACGCTCGGCCACCGGATCTCGGTGATGTGAA 563
Db 1002 GGTCAACGCGGGCGGCGGATGTTGTCGCCGACGCGGACTTGACGCGGAGTTGTCGC 1061
QY 564 CCACGTGCTGCCAGCATGTGCGG---CGTCCATGCTGACCTGGATCGAAACCGGCA 620
Db 1062 CCGGAGGTGGCGGCGCTGTGGCGACCCGCGGCTGGGGCGGATACGACGCGCGC 1121
QY 621 GGAGCGGATCGGAAGTTCGCGCGCGCGCGGCGCGGCGGCTGGAGGCTGGAGGCTCCG 679
Db 1122 CCGCGGGTTCGGGATCCCGACGCGCGCGCGCGGCTGGCGGCGCGGCTGGACATCG 1180

RESULT 11
AD051695/c
ID AD051695 standard; DNA; 32329 BP.
XX AC AD051695;
XX DT
XX 15-JUL-2004 (first entry)
XX Streptomyces cattleya NRRL 8057 thienamycin biosynthetic gene cluster.
KW Beta-lactam synthetase; carbapenem antibiotic; thienamycin;
KW Gram-positive pathogenic bacteria; Gram-negative pathogenic bacteria;
KW infectious disease; therapy; gene; ds.
XX OS Streptomyces cattleya.
XX PH
XX Key Location/Qualifiers
XX CDS complement(2..751)
XX FT /*tag= a
XX FT /product= "Streptomyces cattleya orfX protein"
XX FT /partial
XX FT /note= "No stop codon"
XX CDS 1034..1924
XX FT /*tag= b
XX FT /product= "Streptomyces cattleya orfX1 protein"
XX CDS complement(1940..2719)
XX FT /*tag= c
XX FT /product= "Streptomyces cattleya orfX2 protein"
XX CDS complement(2850..3932)
XX FT /*tag= d
XX FT /product= "Streptomyces cattleya orfZ1 protein"
XX CDS 4117..4755
XX FT /*tag= e
XX FT /product= "Streptomyces cattleya orfZ2 protein"
XX CDS 4945..4936
XX FT /*tag= f
XX FT /product= "Streptomyces cattleya orfZ3 protein"
XX FT /partial
XX FT /note= "No start codon"
XX CDS 6288..7172
XX FT /*tag= g
XX FT /product= "Streptomyces cattleya orf1 protein"
XX CDS complement(7156..8139)
XX FT /*tag= h
XX FT /product= "Streptomyces cattleya orf1b protein"
XX FT /partial
XX FT /note= "No start codon"
XX CDS complement(8136..8927)
XX FT /*tag= i
XX FT /product= "Streptomyces cattleya orf1c protein"
XX CDS 9171..9845
XX FT /*tag= j
XX FT /product= "Streptomyces cattleya orfA protein"
XX CDS complement(9767..11197)
XX FT /*tag= k

```

US2004038250-A1.

26-FEB-2004.

XX

PF 26-FEB-2003; 2003US-00374903.
XX P-PSDB; ADO51696, ADO51697, ADO51698, ADO51699, ADO51700, ADO51701,
XX ADO51702, ADO51703, ADO51704, ADO51705, ADO51706, ADO51707, ADO51708,
XX ADO51709, ADO51710, ADO51711, ADO51712, ADO51713, ADO51714, ADO51715,
XX ADO51716, ADO51717, ADO51718, ADO51719, ADO51720, ADO51721, ADO51722,
XX ADO51723, ADO51724, ADO51725.
XX Isolation, cloning and sequencing of the gene cluster involved in the
PT biosynthesis of the carbapenem thienamycin, useful for production of
PT antibiotics.
XX
PS Claim 2; SEQ ID NO 1; 55pp; English.
XX
CC The present invention relates to the isolation, cloning and sequencing of
CC the gene cluster encoding a beta-lactam synthetase involved in the
CC biosynthesis of the carbapenem antibiotic thienamycin. The invention thus
CC relates to the novel polypeptides and their encoding genes and nucleic
CC acids. The invention is useful for the production of beta-lactam
CC antibiotics and beta-lactam precursors especially thienamycin or a
CC thienamycin precursor which is active against Gram-positive and Gram-
CC negative pathogenic bacteria, both aerobic and anaerobic and is useful in
CC the treatment of infectious diseases. The present sequence is
CC Streptomyces cattleya NRRL 8057 thienamycin biosynthetic gene cluster.
XX
SQ Sequence 32329 BP; 4151 A; 11383 C; 12486 G; 4309 T; 0 U; 0 Other;

Query Match 9.8%; Score 68.2; DB 12; Length 32329;
Best Local Similarity 47.2%; Pred. No. 0.0051;
Matches 307; Conservative 0; Mismatches 338; Indels 6; Gaps 3;

QY 49 CGCAGCGCGGAGCGATCGTGGCGGAGCGCGCTCGGTCTTCGACGAGTACGGGTTCGAG 108
DB |||||
DB CGCAACGGCGCAAGGTGTGGACGCGCGCGCGGAGGAGTTCCGCCAGCGGCTACCGG 3840

QY 109 GCGCCACAGTGGCAGAGATCTCTTCGCGGGCTCGGTTCACCAAGGCGCGATGTACTTC 168
DB |||||
DB GACGCCAGGATCGACGCATCGCGGAGCGCGCGCTCACCGTGGTGGTCTACTCTC 3780

QY 169 CACTTCGCTTCCAGAGAGCTGGCGCGCGGCTGCTGGCGGAGCAGACCTTCGACGTG 228
DB |||||
DB AACTTCCCGGCAAGCGGCGCTTACTTGGCGGCTGCTGGCGGCTTGGCGGAAACGCGTG 3720

QY 229 GCGGTGCGGAAATCCGGCTCAAGGCGCAGGAACTGTAGACCTACCATGCTGGTCCGC 288
DB |||||
DB CCGAGGAGGACCGCGGAGCGGTGCGTCCGCGGCGGAGCGCTTGGCGGCTTCGCC 3660

QY 289 CACGGCATGTGCACGATCCGATCTCTCGCGGCGGCGACGCGGCTCGCACTGGACACAGGG 348
DB |||||
DB CGGGCTTGGGTCTCGCGGCTCGCGCTCGCGCGGCGGAGCGCGCGCGCGCGCGCTC 3600

QY 349 GCGGTGGAC---TTCTCGGAGCGCAACCGCTTTCGGGAGTGGGGGCGACATCTCGGCCGAG 405
DB |||||
DB GCCATGGACCTGATGCCGAGGTGATCACGACGCGCGCGCACCGCGCGGCGCTTCGCCGAG 3540

QY 406 CTCCTCGCGGAGGACAGGAAACCGGGGAGTGTCTTCGCGACGTGAACCGGAAAGACG 465
DB |||||
DB CTGCTGAAGCTCAGCGGATCTCTGCGGCTCGCGCTGGAACGG---CTGGAGCCCGCG 3482

QY 466 GCGGACTTTCATCGTGGCTGCTTTCACCGGCGCTCCAGCGGTCTCCGGGTTCACCTCCGAG 525
DB |||||
DB GACGCCCGCGCGCGATGTCGGGTGCGCGGCGCGGCGGCTCTCACGCGACTGCACGGCA 3422

QY 526 GCGCAGGACCTCGGCGCACCGGATCTCTGGTGAATGTGAAACCACTGTCTGCCAGCATGCTG 585
DB |||||

DB 3421 CGGCCCACTGCGCGCGCGCGCGCGGTTTCTCGACCGCTCGAGCGGTACGGGTCT 3362
QY 586 CCGGCGTCCATCTGACCTGGATCGAAACCGCGGAGAGCGGATCGGAGAGTTCGCGCG 645
DB |||||
DB GCGAGCGCTGCGCGGCTGCGCACTGGA-CGACGACCGCGCGCGCGCGCGCTGCGCG 3303
QY 646 GCGGCGGAGCGCGCGGAGGCTCGGAGGCTCGGAGGCGCGCTCCGACGAG 696
DB |||||
DB CCCGTCACACGCTCGACGAAACCGTGTCTCCCGCGCGCGCGCGCGCGCGCG 3252

RESULT 12
AAQ76213/c
ID AAQ76213 standard; DNA; 12001 BP.
XX
AC AAQ76213;
XX
XX 25-MAR-2003 (revised)
DT 04-AUG-1995 (first entry)
XX
DE HSV L/ST region.
XX
KW HSV; junction-spanning transcript; L/ST; therapy; virucide; ds.
XX
OS Herpes simplex virus type 1.
XX
FH Key Location/Qualifiers
FT TATA_signal 199..204 /tag= a
FT misc_binding 220..225 /tag= b
FT /note= "ICP4 binding site"
FT misc_RNA 229 /tag= c
FT /note= "5' end of the L/STs"
FT misc_RNA 370..372 /tag= d
FT /note= "first codon of ORF-1"
XX
PN W09428156-A1.
XX
PD 08-DEC-1994.
XX
PF 20-MAY-1994; 94WO-US005770.
XX
PR 20-MAY-1993; 93US-00065146.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Schaffer PA, Yeh L;
XX
XX WPI; 1995-022825/03.
XX
XX Herpes Simplex Virus (HSV) specific junction spanning transcript - for
PT inhibiting HSV L/ST synthesis, in the treatment of HSV infection.
XX
PS Disclosure; Page 38-44; 64pp; English.
XX
XX An HSV-specific junction-spanning transcript (L/ST) maps at the 5' end to
CC the b repeat sequences of HSV DNA at approx. 3 and 125 kb, and extends
CC into the c repeat sequences of HSV DNA. Compounds that inhibit HSV L/ST
CC synthesis may be used as anti-HSV virucides. The HSV-1 DNA sequence in
CC the region of the L/STs is given in AAQ76213. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
XX
SQ Sequence 12001 BP; 1568 A; 4557 C; 4315 G; 1561 T; 0 U; 0 Other;
QY
Query Match 9.7%; Score 67.8; DB 2; Length 12001;
Best Local Similarity 45.0%; Pred. No. 0.0064;
Matches 293; Conservative 0; Mismatches 357; Indels 1; Gaps 1;

QY 43 GCGGTCCGACGCGCGGAGCGGATCGTGGGAGCGCGCTCGTCTTCGACGAGTAGCGG 102
DB |||||

Db 4806 GCCGTGCGCGCGCCGACACCGCGGAGCGCGCGGTGCTCCCGCGGTACCGCGCGCGCGG 4747
Qy 103 TTCGAGGCGCGCACAGTGCAGAGATCCTCTCGCGGCGCTCGTCAACCAAGGCGCGATG 162
Db 4746 ATCTCTCGCGCC - TGGGGCGGTGTTCGCGCGCGCGCTCCCGCGCGGCGCGACGAC 4688
Qy 163 TACTTTCACCTTCGCTTCCAGGAAGAGCTGGCGCGGTGTGTGCGCGAGCAGACCCCTG 222
Db 4687 CCGGAGCGCGCGCCACCGCGAGCGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 4628
Qy 223 CAGGTGGCGGTGCGGAATCCGGCTCAAGAGCGAGGAATCGTAGACCTCAACATGCTG 282
Db 4627 CCGGTGGCGGTGCGAGTGCCTGCGCGCTTCGCGCGGATCTCGAGCGGTGCGCGAGGCG 4568
Qy 283 GTGCCCGACGATGTCGACGATCCGATCCTGCGCGGCGGCGAGCGCGGTGCGCACTGGAC 342
Db 4567 TTCGACGGGACCTGCGCGCGCTTCGCGGCGGTGCGCGGCGCGCGCGCGCGCGCGCG 4508
Qy 343 CAGGGGCGGTGGACTTCTCCGACGCCAACCCGTTTGGCGAGTGGGGCGACATCTCGGCC 402
Db 4507 CGGCGGAGGACCGCGCGGCGCGGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4448
Qy 403 CAGCTCTCGGAGGACAGAAACCGGGGAGGTGCTTCCGCACTGTAACCGGAAAG 462
Db 4447 CTGCGCGGTGCTGCGGAGCTGCGGTTCGTGCGCGACGCGCTGTGCTCATGCGCTG 4388
Qy 463 ACCGGGACTTCATCTCGGTGCTTCCACGCGGCTCCAGCGGTCTCCGCGGTCACTCC 522
Db 4387 CCGGGGACCTGCGGTGCGCGGCGCGGACGCGCGCGCGCGCGCGCGCGCGCGCGCG 4328
Qy 523 GACCGCGAGACTCGGCGCACCGGATCTCGGTGATGTGGAACCACTGCTGCGCCAGCATC 582
Db 4327 AGCTTGTGCGCGGCGCGCTGGGCGCGCGCTGCGCGGACCGCGCGCTGCGGACTCC 4268
Qy 583 GTGCGCGGTCTCATGTGACCTGATCGAATCGGAGGAGCGGATCGGGAAGTTCGCG 642
Db 4267 GCGGCGCGCGCGCGCGGACCTGCTGTTTGAACACGAGACTGCGCGCGCGCTGCTGGCG 4208
Qy 643 GCGGCGCGCGCGCGCGGCTGCGGAGGCTTCGAGGCGCGCTCCGAGCGCGCTCCGAC 693
Db 4207 GCGGCGCGCGCGCGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4157

RESULT 13
ID AAS54248 standard; DNA; 1041 BP.
XX AAS54248;
AC AAS54248;
XX
DT 13-FEB-2002 (first entry)
DE Pseudomonas aeruginosa DNA for cellular proliferation protein #379.
KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
XX antibacterial; drug design.
XX Pseudomonas aeruginosa.
OS
XX WO200170955-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US009180.
XX
XX 21-MAR-2000; 2000US-0191078P.
XX 23-MAY-2000; 2000US-0206848P.
XX 26-MAY-2000; 2000US-0207727P.
XX 23-OCT-2000; 2000US-0242578P.
XX 27-NOV-2000; 2000US-0253625P.
XX 22-DEC-2000; 2000US-0257931P.
XX 16-FEB-2001; 2001US-0269308P.
XX
XX (ELIT-) ELITRA PHARM INC.

XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
DR P-PSDB; AAU36389.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Claim 27; SEQ ID NO 7885; 511bp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes,
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/publ/published_pct_sequences
XX
SQ Sequence 1041 BP; 157 A; 347 C; 385 G; 152 T; 0 U; 0 Other;
Query Match 9.6%; Score 67.2; DB 4; Length 1041;
Best Local Similarity 46.9%; Pred. No. 0.0097;
Matches 210; Conservative 0; Mismatches 238; Indels 0; Gaps 0;
Qy 212 AGCAGACCTTGACCTGGCGGTGCGCGGATCCCGGCTCCAAAGCGCGAGGAACCTGGTAGACC 271
Db 575 AGCGCGCATCGAGCTGCGCGGCGCTATGGCGACGAGGTGATGCTGGAGCGCTTCGTGCG 634
Qy 272 TCACATGCTGGTCCCGCACCGCATGCTGCAGCATCCGATCTGCGGGCGGCGACGCGGC 331
Db 635 CCGGCGCGGAGGTCAACCGTCCGGGTGCTCGACGACGACGCGCTGCGGTGGCGGATTC 694
Qy 332 TCGCACTGACCGAGCGCGGTGGACTTCTCCGACGCCAACCCGTTCCGCGAGTGGGGCG 391
Db 695 TCCTCGGCGCGCAGGAGTCTTCGACTAGAGCAACGATACGAGCGCGCGCGGTGCGCG 754
Qy 392 ACATCTGCGCGCCAGCTCTCTGCGGAGGCGACAGGAACGGGGGAGGTGCTTCCGACGTGA 451
Db 755 AGGTGTTTCCCGCGACCTGCGCGCGCGATCGCGCGAGGCCAGCGCTCTGGCGCTGA 814
Qy 452 ACCCGAAAAGACCGCGGACTTCATGCTGGGTGCTTACCGGGGTCCAGGCGGTCTCC 511
Db 815 AGGTACACCGGGCGCTGAAGCTGAGCGGTACAGCGCGACCGACTTCGCGCTCGAGAAC 874
Qy 512 GGGTCACCTCCGACCGCGAGGACCTCGCGCACCGGATCTCGGTGATGTGGAACACGCTGC 571
Db 875 AGGGCGGCTCTGGTGCGCTGGAGGTCAATACCTCGCGGGGATGACCGCCACGAGCTGC 934
Qy 572 TGCCCGAGCATCGTCCCGCGCTCCATGCTGACCTGGATCGAAACCGCGAGGACGCGATCG 631
Db 935 TGCGCGAGCGCGCGCGCGCGCGGATCGGTTTCGCGGAACCTCTCGAGCGGATCTGCC 994
Qy 632 GGAAGTTCG 659
Db 995 GGCCTCGCATCGAGCGCTGCAAGGCGCG 1022
RESULT 14
ID ACA42557 standard; DNA; 1041 BP.

XX ACA42557;
XX 19-JUN-2003 (first entry)
XX Prokaryotic essential gene #24214.
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX Pseudomonas aeruginosa.
XX W0200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX P-PSDB; ABU38687.
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 14; SEQ ID NO 30427; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
XX prokaryotic essential genes. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 1041 BP; 157 A; 347 C; 385 G; 152 T; 0 U; 0 Other;
Query Match 9.6%; Score 67.2; DB 8; Length 1041;

Best Local Similarity 46.9%; Pred. No. 0.0097;
Matches 210; Conservative 0; Mismatches 238; Indels 0; Gaps 0;
QY 212 AGCAGACCTGACAGTGGCGGTGCGGAATCGGGTCCAAAGGCGCAGGAACCTGGTAGACC 271
DB 575 AGCCGGCCATCGAGCTGGCGGCGCTATGGGCACAGGTGATGCTGGAGCGCTTGTGTCG 634
QY 272 TCACCATGCTGGTGCACCGGATCTGCAGATCCGATCTTCGCGGGGCGCACGCGGC 331
DB 635 CCGGCGCGAGGTTCACCGTGGGGTCTCGACGACGAGCGCTGCGGTGGGCGAGATTC 694
QY 332 TCGCACTGACACGAGGGGCGGTGGAATCTTCGACGCCAACCGTTTCGGGAGTGGGGCG 391
DB 695 TCTCGCGGCGCAGGAGGTCTTCGACTACGAGCACAAGTACCAGGCGCGCGGTGCGCG 754
QY 392 ACATCTGCGCCCGAGCTCTTGGCGGAGGCACAGAAACGGGGGAGGTCTTCCCGACGCTGA 451
DB 755 AGGTGTTTCCCGCGGACCTGCGCGCGCGATGCCCGCGAGGCCGCTCTGGCGCTGA 814
QY 452 ACCCGAAAAGACCGGCGACTTCATCGTGGGTGCTTACCGGGCTCCAGGCGGTCTCC 511
DB 815 AGGTACACCGGCGCTGAAGCTGAGCGGCTACAGCGGCACCGACTTCCGCTCGACGAAC 874
QY 512 GGTCACTCCGACCGCCAGGACCTCGGCCACCGGATCTCGGTGATGTGGAACACGTCG 571
DB 875 AGGGCGCGCTCTGTGCTGAGGTCAATACCTGCGCGGCGATGACCGCCACGAGCTGC 934
QY 572 TGCCCAAGCATCTGCGCGGCTCCATGCTGACCTGGATCGAAACCGCGGAGGCGGATCG 631
DB 935 TGCCCGAGCGCGCGGCGCGCGGATCGGTTCGCCGAACCTCTCGAGCGGATCTGCC 994
QY 632 GGAAGTCCGCGCGCGCGCGCGCGCGC 659
DB 995 GGTCTGGCATCGAGCGCTCAAGGGGCGC 1022
RESULT 15
ABD05332
ID ABD05332 standard; DNA; 1107 BP.
XX AC ABD05332;
XX 29-JUL-2004 (first entry)
XX Pseudomonas aeruginosa polynucleotide #3936.
XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
XX antibacterial.
XX Pseudomonas aeruginosa.
XX US6551795-B1.
XX 22-APR-2003.
XX 18-FEB-1999; 99US-00252991.
XX 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX P-PSDB; AB071761.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX Disclosure; SEQ ID NO 3936; 455pp; English.
XX

us-10-049-710a-1.rng

The invention relates to *Pseudomonas aeruginosa* polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a *P. aeruginosa* nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-*P. aeruginosa* drugs, as templates for recombinant production of *P. aeruginosa*-derived peptides or polypeptides, as target components for diagnosis and/or treatment of *P. aeruginosa*-caused infection, and in detection of *P. aeruginosa* sequences or other sequences of *Pseudomonas* species using biopchip technology. Sequences ABD01397-ABD17967 represent *P. aeruginosa* polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html

Sequence 1107 BP; 168 A; 366 C; 401 G; 172 T; 0 U; 0 Other; XX SQ

Query Match 9.4%; Score 66; DB 11; Length 1107;

Best Local Similarity 47.6%; Pred. No. 0.016;

```

Best local similarity: 1.00; 100% identity
Matches 195; Conservative 0; Mismatches 215; Indels 0; Gaps 0;

```

QY 212 AGCAGACCTGCACGCTGGCGGTGCCGGAATCCGGCTCCAAGCGCAGGAAC TGGTAGACC 271

Db 641 AGCCGGCCATCGAGCTGGCCGGCGCTATGGCGACGAGGTGATGCTGGAGCGCTTCGTCTG 700

272 TCACCATGCTGTCGCCACGGCATGCTGCACGATCCGATCCTCGGGGGCAGCGGC 331

701 CCGGCCGAGGTCAACCGTCGGGGTGCTCGACGACAGGCGCTGCCGTGGCGAGATC 760

332 TCGCACTGGACCAAGGGGGCGGTGGACTTCTCCGACGCCAACCCGTTCCGCCAGTGGGGCG 391

761 TCCTCGECGGCCACGGAGGTCTTCGACTACGAGCACAGTACCAGCCGGCGGTGCGG 820

392 ACATCTGGCGCCAGCTCCTGGCGGAGGCA CAGGAA CCGGGGGAGGTGCTTCCGCA CGTGA 451

921 AGGTGTTTCCCGCGGACCTGCCCGCGGGGATCGCCGCGGAGGCCACGCGTCTGGCGCTGA 880

453 ACCCGAAAAAGACCGGCGACTTCATCGTCGGGTGCTTCA CCGGGCTCCAGGCGGTCTCCC 511

981 AGCTACACCCGCCCTGACCTGACGGCTACAGCCGCACCGACTTCGGCCTCGACGAAC 940

571

0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9 2.0 2.1 2.2 2.3 2.4 2.5 2.6 2.7 2.8 2.9 3.0 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 4.0 4.1 4.2 4.3 4.4 4.5 4.6 4.7 4.8 4.9 5.0 5.1 5.2 5.3 5.4 5.5 5.6 5.7 5.8 5.9 6.0 6.1 6.2 6.3 6.4 6.5 6.6 6.7 6.8 6.9 7.0 7.1 7.2 7.3 7.4 7.5 7.6 7.7 7.8 7.9 8.0 8.1 8.2 8.3 8.4 8.5 8.6 8.7 8.8 8.9 9.0 9.1 9.2 9.3 9.4 9.5 9.6 9.7 9.8 9.9 10.0

631

7

.....

Search completed: March 6, 2005, 20:06:59

Job time : 594 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2005, 19:54:40 ; Search time 227 Seconds
(without alignments)
5038.579 Million cell updates/sec

Title: US-10-049-710A-1
Perfect score: 699
Sequence: 1 atggcagtgacacgaacg.....aggccgctccgacagtag 699

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78.2	11.2	4257	2	US-08-690-473-1
2	78.2	11.2	4257	3	US-09-259-821A-1
3	78.2	11.2	4257	3	US-08-843-659-1
4	78.2	11.2	4257	4	US-09-825-288A-1
5	67.8	9.7	12001	1	US-08-458-568A-11
6	66	9.4	1107	4	US-09-252-991A-3936
7	66	9.4	1134	4	US-09-252-991A-3908
8	66	9.4	1326	4	US-09-252-991A-4002
9	65	9.3	77536	4	US-09-410-551B-1
10	65	9.3	77536	4	US-09-940-316B-1
11	62.8	9.0	4403765	3	US-09-103-840A-2
12	62.8	9.0	4411529	3	US-09-103-840A-1
13	62	8.9	630	4	US-09-302-540-4889
14	62	8.9	28493	4	US-09-302-540-1241
15	60.2	8.6	23673	3	US-09-773-816-1
16	59.6	8.5	2721	6	5215881-2
17	59.6	8.5	2721	6	5215881-2
18	58.4	8.4	1536	4	US-09-302-540-8768
19	58.4	8.4	9185	4	US-09-302-540-931
20	58.2	8.3	3957	4	US-10-237-551-193
21	58.2	8.3	154746	4	US-09-827-688-8
22	58.2	8.3	154746	4	US-09-827-688-8
23	58	8.3	8438	1	US-07-945-283-1
24	58	8.3	15872	3	US-09-105-537-1
25	58	8.3	15872	4	US-09-091-609-1
26	58	8.3	15872	4	US-09-091-609-3
27	57.2	8.2	1626	3	US-09-158-767-14

28	57.2	8.2	1626	4	US-09-713-794-14	Sequence 14, Appl
29	57.2	8.2	2181	3	US-09-158-767-10	Sequence 10, Appl
30	57.2	8.2	2181	4	US-09-713-794-10	Sequence 10, Appl
31	57	8.2	34662	4	US-09-902-540-1261	Sequence 1261, Ap
32	56.8	8.1	717	4	US-09-252-991A-11242	Sequence 11242, A
33	56.8	8.1	1632	4	US-09-252-991A-11015	Sequence 11015, A
34	56	8.0	1197	4	US-10-151-832-2	Sequence 2, Appli
35	56	8.0	1227	4	US-09-252-991A-16560	Sequence 16560, A
36	56	8.0	1308	4	US-10-151-832-1	Sequence 1, Appli
37	56	8.0	1368	4	US-09-252-991A-16040	Sequence 16040, A
38	56	8.0	2787	3	US-09-105-537-40	Sequence 40, Appl
39	56	8.0	5970	3	US-09-320-878-21	Sequence 21, Appl
40	56	8.0	5970	4	US-09-141-908-11	Sequence 11, Appl
41	56	8.0	5970	4	US-09-657-440-21	Sequence 21, Appl
42	55.6	8.0	378	4	US-09-902-540-2074	Sequence 2074, Ap
43	55.6	8.0	2028	4	US-09-252-991A-13079	Sequence 13079, A
44	55.6	8.0	3525	4	US-09-252-991A-12708	Sequence 12708, A
45	55.6	8.0	3906	4	US-09-252-991A-13251	Sequence 13251, A

ALIGNMENTS

RESULT 1
US-08-690-473-1
; Sequence 1, Application US/08690473
; Patent No. 5876923
; GENERAL INFORMATION:
; APPLICANT: Leopardi, Rosario
; APPLICANT: Roizman, Bernard
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 AS AN
; TITLE OF INVENTION: INHIBITOR OF APOPTOSIS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/690,473
; FILING DATE: 26-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: ARCD:239
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4257 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-690-473-1

Query Match 11.2%; Score 78.2; DB 2; Length 4257;
Best Local Similarity 45.0%; Pred. No. 1.3e-07;
Matches 293; Conservative 0; Mismatches 358; Indels 0; Gaps 0;
QY 43 GCGTCGCGCAGCGGCGGATCGTGGGGGAGCGGCTCGGTCTTCGACGAGTACGGG 102
Db 1867 GCGTCGCGCGCGCGCGGCGGAGCGGCGGTCGCGGTCGCGCGCGGGG 1926
QY 103 TTCGAGCGCGCCACAGTGGCAGAGATCTCTCGCGGGCCTCGGTCCACAGGGCGGATG 162

```

Db      1927  ATCTCGCCGCGCTTCAAAGAAAGAGTGGCCCGCGCGCTGTCCCGCGCGCCCGCTCCCGCGGGGGGCGACGAC 1986
Qy      163   TACTTCCACTTCGCTTCAAAGAAAGAGTGGCCCGCGCGTGTGGCCGAGCAGACCCCTG 222
Db      1987  CCCGACGCCCGCCGACGCGCGAGCGACGACGACGCCGGGCGCCGCGCCAGCGCGC 2046
Qy      223   CAGGTGGCGGTGCGCGAATCCGGCTCAAGGCGCAGAACTGTGTAGACCTCAACATGCTG 282
Db      2047  CGCGTGGCCGTGCAAGTGCCTGGCGCGCTCGCGGGGATCTGTGAGGCGCTGGCGAGGGC 2106
Qy      283   GTGCCCCACGGCATGCTGCAGATCCGATCTTGGGGGGGGGACGCGGCTGCACTGGAC 342
Db      2107  TTGACCGGCGACCTGGCGCGCGCTCCCGGGGCTGGCGGGGCGCGGCGCCGACGCCCGCG 2166
Qy      343   CAGGGGCGGTGGACTTCTCCGACGCCCAACCCCGTTCGGCGAGTGGGCGGACATCTGCGCC 402
Db      2167  CGGCGGAGGACCCCGGGGCCCGCTTCCCGCGCGCGCGGACGCGCGAGCGCGCCCGC 2226
Qy      403   CAGCTCTGCGGAGGACAGGACAGGAAACGGGGGAGTGTTCGCGACGTGAACCCGAAAAAG 462
Db      2227  CTGCGCGGTGCTGCGCGAGCTGCGGTTCGTGCGGACGCGTGTGTCTCATGCGCGCTG 2286
Qy      463   ACCGGGACTTCATCGTCCGCTGCTTACCGGGCTCCAGGGGTCTCCGGGTCACTCC 522
Db      2287  CGCGGGACCTGCGGTGGCGCGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGT 2346
Qy      523   GACCGCGCAGGACCTCGGCCACCGGATCTCGGTGATGTGAAACCAACGTCGTGCCCCAGCATC 582
Db      2347  AGCTGTGTCGGGGGCCCTGGGCCCGCGCTGCCGCGGACCCGCGCGCTGCCGAGCTCC 2406
Qy      583   GTGCGCGCTCATGTGACTGATCGAATCGAAACCGCGGAGAGCGGATCGGGAAAGTCGCG 642
Db      2407  GCGCGCGCGCCGCGCGGACCTGCTGTTTGAACCCAGAGCCTGCGGCCCTGCTGGCG 2466
Qy      643   GCGCGCGCGAGCGCGCGAGGTGCGAGGCGCTCGAGGCGCGCTCCGAGGCGCGCTCCGAC 693
Db      2467  GCGCGCGCAGGACCGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 2517

RESULT 2
US-09-259-821A-1
; Sequence 1, Application US/09259821A
; Patent No. 6210926
; GENERAL INFORMATION:
; APPLICANT: LEOPARDI, ROSARIO
; APPLICANT: ROIZMAN, BERNARD
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS
; FILE REFERENCE: ARCD:317
; CURRENT APPLICATION NUMBER: US/09/259,821A
; CURRENT FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 08/690,473
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ. ID NO 1
; LENGTH: 4257
; TYPE: DNA
; ORGANISM: HERPES VIRUS, TYPE 1
US-09-259-821A-1

```

Qy	163	TACTTCACATTTCGCTTCCAAAGGAAGAGCTGGCCCGCGCGTGTCTGCGCGAGACCCCTG	222
Db	1987	CCCGACGCCGCCCGCCACACGCCGACCGCCAGCAGCAGCAGCGCGGGCGCGCGCCAGGCGCGC	2046
Qy	223	CAGCTGGCGGTGGCGGAAATCCGGCTCCAGGCGCAGGAACTGTAGACCTCACCATTGCTG	282
Db	2047	CGCGTGGCGGTGCGAGTGCCTGGCGCGCTTGC CGCGGGATCTCTGAGGCGCTGGCGCGAGGC	2106
Qy	283	GTGCCCCACACGGCATGTGTCACGATCCGATCTCTTCGGGCGGGGCACAGCGGGTTCGCATCGAC	342
Db	2107	TTTCAGCGGGACCTGGCGGCCGTCCTCCGGGCTGGCGGGGCCCGCGCCCGCCAGCCCCCG	2166
Qy	343	CAGGGGGCGGTGACATTCCTCCGACGCCAACCGTTCGGCGAGTGGGGCGACATCTGGGCC	402
Db	2167	CGGCCCGAGGAGCCCGCGGGCCCCCGTTCCCCCGCCCGCCGACGCCGACGCGCGCCCCGC	2226
Qy	403	CAGCTCTGTCGGGAGGACACAGAAACGGGGGAGGTCTTCCGCACTGAAACCCGAAAAAG	462
Db	2227	CTGGCGCGTGGCTGTCGCGAGCTGCGGTTCTGTGCGGACGCGCTGTGTCTATGCGCCTG	2286
Qy	463	ACGGCGACATTCAATCTGGGTGCTTACCGGGCTCCAGCGGTCTCCCGGGTCACTCC	522
Db	2287	CGCGGGGACCTGCGCGTGGCGGGCGCAGCAGCGCCCGTGGCCCGCTGGCGCGCGTG	2346
Qy	523	GACCGCCAGGACCTCGGCCACCGGATCTCGGTGATGTGGAACCACTGTGCCACGATC	582
Db	2347	AGCTGTGTGCGCGGGCCCTTGGGCCCGCTGCCCGGACCCCGGCTTCCGAGCTCC	2406
Qy	583	GTCCGGGTCCATGCTGACCTGGATCGAAACCGGCGAGAGCGGATCGGGAAGTFCGCG	642
Db	2407	GCGGCGCGCGCGCCCGGACCTGCTGTTTGAACAACAGAGCTGCGCGCCCTGCTGGCG	2466
Qy	643	CGCGCGCGCGAGCCCGCGAGGCTCGGAGGCTCCGAGGCGCGCTCCGAC	693
Db	2467	CGCGCGCGCAGCGCACCGGACCGCCCGCAGCGCTGGCGCGCCCGCC	2517

RESULT 3

US-08-843-659-1

; Sequence 1, Application US/08843659
; Patent No. 6218103
; GENERAL INFORMATION:
; APPLICANT: Leopardi, Roario
; APPLICANT: Roizman, Bernard
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS US3 AND ICP4 AS
; TITLE OF INVENTION: INHIBITORS OF APOPTOSIS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/843,659
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: ARGB:519
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4257 base pairs

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-843-659-1

Query Match      11.2%; Score 78.2; DB 3; Length 4257;
Best Local Similarity 45.0%; Pred. No. 1.3e-07;
Matches 293; Conservative 0; Mismatches 358; Indels 0; Gaps 0;

QY 43 GCGCTCGGACGCGGAGCGATCGTGGGAGCGCGCTCGTCTTCACGAGTACGGG 102
Db 1867 GCGCTCGGCGCGCGCCACCGGGGAGCGCGGCTCGCGCGGTACGCGCGCGGG 1926
QY 103 TTCGAGGCGCGCACATGCGAGATCTCTTCGCGGGCTCGGTACCAAGGCGCGATG 162
Db 1927 ATCTCGCGCGCGCTGGGGGCGGTGTCTCGCGCGCGCTCTCCCGCGGGGCGACGAC 1986
QY 163 TACTTCCACTTCGTTCCAAAGAGAGCTGGCGCGGCGTGTGCGCGAGACGCTG 222
Db 1987 CCGGACGCGCGCGCCACCGCGACGCGGAGCGCGCGCGCGCGCGCGCGCGCG 2046
QY 223 CACGTGGCGGTGCGGAAATCCGGCTCCAAAGGCGCAGGAACTGTGTAGACCTCACCATGCTG 282
Db 2047 CGCGTGGCGGTGCGAGTGCCTGGCGCGCTGCGCGGGATCTTGAGGCGCTGGCGGAGGC 2106
QY 283 GTGCGCCACGCGATGTCAGATCCGATCTTCGCGGGCGGGACGCGGCTTCGACATGAGAC 342
Db 2107 TTCGACGCGGACCTGGCGCGCTTCCCGGGGCTGGCGGGGCGCGCGCGCGCGCG 2166
QY 343 CAGGGGCGGTGAGCTTCTCCGACGCGCAACCGCTTCGGCGAGTGGGGCGACATCTCGGC 402
Db 2167 CGCGCGGAGGAGCGCGCGCGCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCG 2226
QY 403 CAGCTCTCTGGCGAGGACAGGAAACGGGGGAGGTCTTCCGACGCTGAACCGGAAAG 462
Db 2227 CTGGCGCGGTGCTGCGCGAGCTGCGTTCGTGCGGAGCGCGTGTGTCTATGCGCTG 2286
QY 463 ACCGGGACTTCATCTCGGTGCTTTCACGGGCTCCAGCGGCTTCCGGGTTCACCTCC 522
Db 2287 CGCGGGGACCTGCGGTGCGCGGCGGACGCGGCGCGCTGGCGCGCGCTGGCGCGCTG 2346
QY 523 GACCGCGAGACTCTGGCGCACCGGATCTCGGTGATGTGAACACGCTGTGCCGACATC 582
Db 2347 AGCTTGTGCGCGGGCGCTTGGCGCGCGCTTGGCGCGCGCGCGCGCGCGCGCTCC 2406
QY 583 GTGCGCGCTTCATCTGACCTGATTCGAAACCGCGGAGGCGGATCGGGAAGTCCGCG 642
Db 2407 GCGGCGCGCGCGCGCGCGGACCTGTCTTTGACAAACAGAGCTTGGCGCGCGCTGCTGGCG 2466
QY 643 GCGGCGCGCGCGCGCGCGCGCGGAGCTCGGAGGCTTCGAGGCGCGCTTCGAC 693
Db 2467 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2517
```

```
RESULT 4
US-09-825-288A-1
; Sequence 1, Application US/09825288A
; Patent No. 6723511
; GENERAL INFORMATION:
; APPLICANT: LEOPARDI, ROSARIO
; APPLICANT: ROIZMAN, BERNARD
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS
; FILE REFERENCE: ARCD:317USC1
; CURRENT APPLICATION NUMBER: US/09/825,288A
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 09/259,821
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 08/690,473
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4257
```

```
; TYPE: DNA
; ORGANISM: HERPES VIRUS, TYPE 1
US-09-825-288A-1

Query Match      11.2%; Score 78.2; DB 4; Length 4257;
Best Local Similarity 45.0%; Pred. No. 1.3e-07;
Matches 293; Conservative 0; Mismatches 358; Indels 0; Gaps 0;

QY 43 GCGCTCGGACGCGGAGCGATCGTGGGAGCGCGCTCGTCTTCACGAGTACGGG 102
Db 1867 GCGCTCGGCGCGCGCCACCGGGGAGCGCGGCTCGCGCGGTACGCGCGCGGG 1926
QY 103 TTCGAGGCGCGCACATGCGAGATCTCTTCGCGGGCTCGGTACCAAGGCGCGATG 162
Db 1927 ATCTCGCGCGCGCTGGGGGCGGTGTCTCGCGCGCGCTCTCCCGCGGGGCGACGAC 1986
QY 163 TACTTCCACTTCGTTCCAAAGAGAGCTGGCGCGGCGTGTGCGCGAGACGCTG 222
Db 1987 CCGGACGCGCGCGCCACCGCGACGCGGAGCGCGCGCGCGCGCGCGCGCGCG 2046
QY 223 CACGTGGCGGTGCGGAAATCCGGCTCCAAAGGCGCAGGAACTGTGTAGACCTCACCATGCTG 282
Db 2047 CGCGTGGCGGTGCGAGTGCCTGGCGCGCTGCGCGGGATCTTGAGGCGCTGGCGGAGGC 2106
QY 283 GTGCGCCACGCGATGTCAGATCCGATCTTCGCGGGCGGGACGCGGCTTCGACATGAGAC 342
Db 2107 TTCGACGCGGACCTGGCGCGCTTCCCGGGGCTGGCGGGGCGCGCGCGCGCGCG 2166
QY 343 CAGGGGCGGTGAGCTTCTCCGACGCGCAACCGCTTCGGCGAGTGGGGCGACATCTCGGC 402
Db 2167 CGCGCGGAGGAGCGCGCGCGCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCG 2226
QY 403 CAGCTCTCTGGCGAGGACAGGAAACGGGGGAGGTCTTCCGACGCTGAACCGGAAAG 462
Db 2227 CTGGCGCGGTGCTGCGCGAGCTGCGTTCGTGCGGAGCGCGTGTGTCTATGCGCTG 2286
QY 463 ACCGGGACTTCATCTCGGTGCTTTCACGGGCTCCAGCGGCTTCCGGGTTCACCTCC 522
Db 2287 CGCGGGGACCTGCGGTGCGCGGCGGACGCGGCGCGCTGGCGCGCGCTGGCGCGCTG 2346
QY 523 GACCGCGAGACTCTGGCGCACCGGATCTCGGTGATGTGAACACGCTGTGCCGACATC 582
Db 2347 AGCTTGTGCGCGGGCGCTTGGCGCGCGCTTGGCGCGCGCGCGCGCGCGCGCTCC 2406
QY 583 GTGCGCGCTTCATCTGACCTGATTCGAAACCGCGGAGGCGGATCGGGAAGTCCGCG 642
Db 2407 GCGGCGCGCGCGCGCGCGGACCTGTCTTTGACAAACAGAGCTTGGCGCGCGCTGCTGGCG 2466
QY 643 GCGGCGCGCGCGCGCGCGCGCGGAGCTCGGAGGCTTCGAGGCGCGCTTCGAC 693
Db 2467 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2517
```

```
RESULT 5
US-08-458-568A-11/c
; Sequence 11, Application US/08458568A
; Patent No. 5821339
; GENERAL INFORMATION:
; APPLICANT: Schaffer, Priscilla A.
; APPLICANT: Yeh, Lily
; TITLE OF INVENTION: Compositions and Methods for Treatment of Herpesvirus
; TITLE OF INVENTION: Infections
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5821339ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,568A
; FILING DATE: 02-JUNE-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/065,146
; FILING DATE: 05-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn R.
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: DFIC-00029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Herpes simplex virus
; STRAIN: Herpes Simplex Virus Type 1
; US-08-458-568A-11

```

Query Match	9.7%	Score 67.8	DB 1	Length 12001
Best Local Similarity	45.0%	Pred. No. 2.4e-05		
Matches 293	Conservative 0	Mismatches 357	Indels 1	Gaps 1
QY	43	GCCTTCGACGCGGCGAGGCGATCTGTCGGGGCAGCCGCTTCGGTCTTCGACGAGTACGGG	102	
DB	4806	GCCGTGCGCGCGCCGACCCGGCGAGCGCGGTGCTCCGCGGGTACGCGCGCGGGG	4747	
QY	103	TTGAGGCGCCACAGTGGCAGAGATCTCTCGGGGCTTCGGTTCACAAAGGGCGGATG	162	
DB	4746	ATCTCTGCGCCCC-TGGGGCGGCTGTCCGCGCGCCCGCTCCCGCGGGGGCGACGAC	4688	
QY	163	TACTTCCACTTCGCTTCCAAAGGAAGACTGGCCCGCGCGTGTGGCCGAGCAGACCCCTG	222	
DB	4687	CCGACGCGCGCCGACGCGCGACGACGACGCGCGGCGCCGCGCCAGGCGGC	4628	
QY	223	CAGTGGCGGTGCCGGAATCTCGGCTCCAGGGCGAGGAACTGTGTAGACTCAACATGCTG	282	
DB	4627	CGGCTGGCCGTGAGTGCCTTGGCGCGCTGCGCGGGATCTTGGAGGCGCTGGCGAGGGC	4568	
QY	283	GTGCCCCACGGCATGCTGCACGATCCGATCCTCGGGCGGGCAGCGGGCTCGCACATGGAC	342	
DB	4567	TTTGA CGGCGA CTGGCGCCGTCTCCGGGGCTGGCCGGGGCCCGGCCCGCAGCCCCCG	4508	
QY	343	CAGGGGCGGTGGACTTCTCCGACGCGCAACCCGTTCCGGCGAGTGGGGCGACATCTGCGCC	402	
DB	4507	CGGCGGAGGGAACCGCGGGCGCCGCTTCCCGCGCGCGCGCGACGCGCGACGCGCCCGC	4448	
QY	403	CAGCTCTGCGGAGGACACAGGAACGGGGGAGGTGCTTCGGCACGTGAACCCGAAAAAG	462	
DB	4447	CTGGCGGTGTGTCGCGAGCTGCGGTTCTGTGGCGACGCGCTGTGTGCTCATCGGCGTG	4388	
QY	463	ACCGGCACTTCATCTGCTGGCTCTTCCACGGGGTCCAGGCGGTCTCCGGGTCACTCC	522	
DB	4387	CGCGGGACCTGCGCGTGGCGGGCAGCGAAGCCGCGCGCGTGGCGCGCGTGCGCGCGTG	4328	
QY	523	GACGCCAGGACCTCGGCCACCGGATCTCGGTGATGTGGAAACACGCTGTGCCCGCAGATC	582	
DB	4327	AGCCTGTGTCGCGGGGCCCTTGGGCCCGCGCTGCGCGGGACCCGCGCCCTTGCAGAGCTCC	4268	
QY	583	GTCCGCGCGTCCATGTCTGACTTGATCGAACCAGCGCGAGGCGGATTCGGGAAGGTGCGG	642	

```

Db      4267  GCGGCGCCCGCCGCGCGGACCTTGCTTTTGA CAACAGAGACCTGCGCCCGCTGCTGGCG 4208
Qy      643  GCGGCGGCGGAGGCGCGCGGAGGCTCGGAGGCGCTCGAGAGCGCGCTCCGAC 693
Db      4207  GCGGCGGCGGAGGCGACCGGACGCGCGCGACGCGCTGCGGCGCGCGCGCGCC 4157

RESULT 6
US-09-252-991A-3936
; Sequence 3936, Application US/09252991A
; Patent NO. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUD
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3936
; LENGTH: 1107
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3936

```

RESULT 7
US-09-252-991A-3908
? Sequence 3908, Application US/09252991A
? Patent No. 6551795
? GENERAL INFORMATION:
? APPLICANT: Marc J. Rubenfield et al.
? TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
? TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
? FILE REFERENCES: 107196.136
? CURRENT APPLICATION NUMBER: US/09/252,991A
? CURRENT FILING DATE: 1999-02-18

```
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1908
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3908

Query Match          9.4%; Score 66; DB 4; Length 1134;
Best Local Similarity 47.6%; Pred. No. 4.1e-05;
Matches 195; Conserved 0; Mismatches 215; Indels 0; Gaps 0;

QY 212 AGCAGACCTTGCACGTGGCGGTGCCGAATCCGGCTCCAAAGCGCAGGAACTGGTAGACC 271
DB 525 AGCGGCCCATCGAGCTGGCGGCGCTATGGCGACGAGGTGATGCTGGAGCGCTTCGTGCG 584
QY 272 TCACCATGCTGTGCGCCACGGATGCTGCAGATCCGATCCTGCGGGCGGGCAGCGCGC 331
DB 585 CGGGCGCGAGGTCAACGCTGGGGTGTCTGACGACACGAGCGCTGCCGGTGGCGGAGATTC 644
QY 332 TCGCACTGACAGGCGGGGTGGACTTCTCCAGCCCAACCGTTCCGGCGAGTGGGGG 391
DB 645 TCTCGCGCGCAGGAGGTCTTCGACTAGACACAAGTACCAGCGCGCGCGGTGCGCG 704
QY 392 ACATCTGCGCCCGAGCTCTTGGCGGAGGACAGAAACGGGGGAGGTGCTTCCGCGACGTGA 451
DB 705 AGGTGTTTCCCGCGACCTGCCGCGCGGATCGCGCGGAGGCCAGCGCTCTGGCGCTGA 764
QY 452 ACCGGAAGAACCGCGAGCTTCATCGTGGCTGTTCACCGGGCTCCAGGCGGTCTCCC 511
DB 765 AGGTACACCGGGCGCTGAAAGCTGAGCGGTACAGCGCGCACCGACTTCCGCGCTCGAGAAC 824
QY 512 GGTCACTCCGACCGCAGGACCTCGGCCACCGATCTCGGTGATGTGGAACCAAGTGC 571
DB 825 AGGGCGCGTCTGTGCTGAGGTCAATACCTCGCGGCGATGACCGCCACAGCGCTGC 884
QY 572 TGCCAGCATCTGCGCGGTCCATCTGACCTGATCGATCGAAACCGCGGAG 621
DB 885 TGCCGAGCGCGCGCGCGCGGATCGGTTTCGCCGAATCTCTCGAG 934

RESULT 8
US-09-252-991A-4002/c
; Sequence 4002, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4002
; LENGTH: 1326
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4002

Query Match          9.4%; Score 66; DB 4; Length 1326;
Best Local Similarity 47.6%; Pred. No. 4.2e-05;
Matches 195; Conserved 0; Mismatches 215; Indels 0; Gaps 0;

QY 212 AGCAGACCTTGCACGTGGCGGTGCCGAATCCGGCTCCAAAGCGCAGGAACTGGTAGACC 271
DB 506 AGCGGCCCATCGAGCTGGCGGCGCTATGGCGACGAGGTGATGCTGGAGCGCTTCGTGCG 447
```

```
QY 272 TCACCATGCTGTGCGCCACGGCATGCTGCAGATCCGATCTGCGGGGCGGACGCGGC 331
DB 446 CGGGCGCGAGGTCAACGCTGGGGTGTCTCGACGACAGCGCGTGGCGGAGATTC 387
QY 332 TCGCACTGACAGGCGGGGTGGACTTCTCCAGCCCAACCGTTCCGGCGAGTGGGGG 391
DB 386 TCTCGCGCGCAGGAGGTCTTCGACTAGACACAAGTACCAGCGCGCGGTGCGCG 327
QY 392 ACATCTGCGCCCGAGCTCTTGGCGGAGGACAGAAACGGGGGAGGTGCTTCCGCGACGTGA 451
DB 326 AGGTGTTTCCCGCGACCTGCCGCGCGGATCGCGCGGAGGCCAGCGCTCTGGCGCTGA 267
QY 452 ACCGGAAGAACCGCGAGCTTCATCGTGGCTGTTCACCGGGCTCCAGGCGGTCTCCC 511
DB 266 AGGTACACCGGGCGCTGAAAGCTGAGCGGTACAGCGCGCACCGACTTCCGCGCTCGAGAAC 207
QY 512 GGTCACTCCGACCGCAGGACCTCGGCCACCGATCTCGGTGATGTGGAACCAAGTGC 571
DB 206 AGGGCGCGTCTGTGCTGAGGTCAATACCTCGCGGCGATGACCGCCACAGCGCTGC 147
QY 572 TGCCAGCATCTGCGCGGTCCATCTGACCTGATCGAAACCGCGGAG 621
DB 146 TGCCGAGCGCGCGCGCGGATCGGTTTCGCCGAATCTCTCGAG 97
```

```
RESULT 9
US-09-410-551B-1
; Sequence 1, Application US/09410551B
; Patent No. 6503737
; GENERAL INFORMATION:
; APPLICANT: KOSAN BIOSCIENCES, Inc.
; APPLICANT: REEVES, CHRISTOPHER
; APPLICANT: CHU, DANIEL
; APPLICANT: KHOSLA, CHAITAN
; APPLICANT: SANTI, DANIEL
; APPLICANT: WU, KAI
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
; FILE REFERENCE: 30062-20026.00
; CURRENT APPLICATION NUMBER: US/09/410,551B
; CURRENT FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/139,650
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: US 60/123,810
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/102,748
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 77536
; TYPE: DNA
; ORGANISM: Streptomyces hygroscopicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52275) ... (71465)
US-09-410-551B-1
```

```
Query Match          9.3%; Score 65; DB 4; Length 77536;
Best Local Similarity 44.7%; Pred. No. 0.00012;
Matches 294; Conserved 0; Mismatches 360; Indels 3; Gaps 1;

QY 41 GGGCGGTTCGACAGCGCGGAGCGGATCGTGGCGGCGAGCGCCCTCGGTCTTCGACGAGTAGC 100
DB 71742 GGGACTTCACACTGCGCGCGCGCGCAGCGGAGGACTTCGTCGCGGCGCGCGGAG 71801
QY 101 GGTTCGAGGCGCGCCACAGTGGCAGAGATCCTTCGCGGGGCGCTCGGTTCACAGGGGCGCA 160
DB 71802 CCTGCTGACGACATCGAGGCGCGGAGCGCGGAGCCCGGACCGACCTCATCTCCCGGTACGCCA 71861
QY 161 TGTACTTCCACTTCGCTTCCAGGAGAGAGCTGCGCGCGCGGTGCTGCGCGGAGACGCC 220
```

Db 71862 AGCGGCTGCCCTCCCTCGTCAATCAACGCGCTGTACCGGGCTCACCCCTGAGGGGGCGCG 71921

QY 221 TGCACTGGGGGTGCGCGGAATCCGGCTCCAAAGCGCGAGAACTGGTAGACTCACCATGC 280

Db 71922 TGTGGAGGACCGATGCGGACATCACCGGCTCGGCCGATCTGGACAGGCTCAAGACG 71981

QY 281 TGTGCGCCACGGCATGCTGACGATCCGATCTCGGGGGGCGACGCGGCTCGCACTGG 340

Db 71982 TGACCGACGACTTCTTTCGGGACGCGCTGCGGCTGGTCCGGCGAAGCGGTGACGAGCGG 72041

QY 341 ACCAGGGGGGTGGAGTCTTCGAGCCCAACCCGTTCCGGCGAGTGGGGGCGACATCTGG 400

Db 72042 GCGAGGACTGCTGCAACCGGCTGGCTCGGCGACGAGCGGAGATCTCGCTCAGCGAG 72101

QY 401 CCAGCTCTTCGGCGGAGGACAGAAACGGGGGAGGTGCTTCGCGACGTAACCCGAAAA 460

Db 72102 ACAGGCGAGGCGGTGTCGCGACGCTGTTCCCGGCGACGACTCGGTGCGACGAG 72161

QY 461 AGACGGGACATTCGTCGGTGTCTTACCGGGGCTCAGGCGGTCTCCCGGGTCACTT 520

Db 72162 TGTGCGGCTACTGCTCTACGCACTGCTCAGCCACCGCGAGCAGGCGGCGCTCGCG 72221

QY 521 CGAGCGCGGAGGCGCGGAGGCTCGGAGGCTCGGAGGCGGCTCGGAGGCGGCTCGGAGG 580

Db 72222 CGGCGCGGAGTGTGCGAACGCGGCTCGAGGAT---GCTCCGTTTCTGCGCGTCA 72278

QY 581 TCGTCCGCGCTCCATGCTGACCTCGATCGAAACCGGCGAGGAGCGGATCGGAAAGTTCG 640

Db 72279 ACCAGATGGCGTACCGCGGCTCTGTGCGAGGACGTGATGTCGGGCGGTGCGCATCC 72338

QY 641 CGGCGCGGCGGAGGCGCGGAGGCTCGGAGGCGGCTCGGAGGCGGCTCGGAGGCGGCT 697

Db 72339 GTGCGGCGGACAACTGATCCCGCTCTACTCGACGCGCAACCGCGACCCCGAGGTGT 72395

RESULT 10

US-09-940-316B-1

; Sequence 1, Application US/09940316B

; Patent No. 6759536

; GENERAL INFORMATION:

; APPLICANT: KOSAN BIOSCIENCES, Inc.

; APPLICANT: REEVES, CHRISTOPHER

; APPLICANT: CHU, DANIEL

; APPLICANT: KHOSLA, CHAITAN

; APPLICANT: SANTI, DANIEL

; APPLICANT: WU, KAI

; TITLE OF INVENTION: POLYKETIDES ENCODING THE FKBA GENE OF THE FK-520 POLYKETIDE SYNTH

; FILE REFERENCE: 30062-20026.11

; CURRENT APPLICATION NUMBER: US/09/940,316B

; PRIOR FILING DATE: 2001-08-27

; PRIOR APPLICATION NUMBER: 09/410,551

; PRIOR FILING DATE: 1999-10-01

; PRIOR APPLICATION NUMBER: US 60/139,650

; PRIOR FILING DATE: 1999-06-17

; PRIOR APPLICATION NUMBER: US 60/123,810

; PRIOR FILING DATE: 1999-03-11

; PRIOR APPLICATION NUMBER: US 60/102,748

; PRIOR FILING DATE: 1998-10-02

; NUMBER OF SEQ ID NOS: 72

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 77536

; TYPE: DNA

; ORGANISM: Streptomyces hygroscopicus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (52275)...(71465)

US-09-940-316B-1

Query Match 9.3%; Score 65; DB 4; Length 77536;
Best Local Similarity 44.7%; Pred. No. 0.00012;
Matches 294; Conservative 0; Mismatches 360; Indels 3; Gaps 1;

QY 41 GGGCGCTCCGCAACGCGGAGGCGATCGTGGGCGAGCGCCCTCGGTCTTCGACGAGTACG 100

Db 71742 GGGACTTACACTGCGCGGCGCGAGCGGAGACTTCTGTCGCCGAGGCGCGCGAG 71801

QY 101 GGTTCGAGGCGCGCCACAGTGGCAGAGATCTCTTCGCGGGCTTCGGTCAACAAGGGCGCA 160

Db 71802 CCGTCTGGACGACATCGAGGCGCGGACCGCGCACGACCTCATCCCGGGTACGCCA 71861

QY 161 TGTACTTCCACTTCGCTTCCAAAGGAAGTGGCGCGGGCGTGTGGCGGAGCAGACCC 220

Db 71862 AGCGGCTGCCCTCCCTCGTCAACAGCGCTGTACGGGCTCACCCCTGAGGAGGGGCGG 71921

QY 221 TGCACTGTCGGTGGCGGAATCCGGCTCCAAAGCGCGAGAACTGGTAGACTCACCATGC 280

Db 71922 TGTGAGGCGAGATGCGCGACATCACCGGCTCGGCGATCTGACAGCGTCAAGACG 71981

QY 281 TGTGCGCCACGGCATGCTGACGATCCGATCTTCGCGGGGCGGACGCGGCTCGCACTGG 340

Db 71982 TGACCGACGACTTCTTTCGGGCGACGCGCTCGGCTGGTCCCGCGAAGCGGTGACGAGCGG 72041

QY 341 ACCAGGGGGGTGGACTTCTCCGAGCCCAACCCGTTCCGGCGAGTGGGGGCGACATCTCG 400

Db 72042 GCGAGGACTGCTGCAACCGGCTGGCTCGGCGACGAGCGGAGATCTCGCTCAGCGAG 72101

QY 401 CCAGCTCTTCGGCGGAGGCGACAGGAAACGGGGGAGGTGCTTCGCGACGTAACCCGAAAA 460

Db 72102 ACAGGCGAGGCGGTGTCGCGACGCTGCTGTCGCCGCGCACGACTCGGTGCGACGAG 72161

QY 461 AGACGGGAGCTTCACTGTCGGCTGCTTCAACCGGGCTCAGGCGGTCTCCCGGGTCACTT 520

Db 72162 TGTGCGGCTACTGCTCTACGCACTGCTCAGCACCCCGAGCAGGCGGCGCTCGCG 72221

QY 521 CGAGCGCGGAGGCGGAGGCTCGGCGGAGGCTCGGAGGCTGGAACCAACGTCGTCGCGAG 580

Db 72222 CGGCGCGGAGTGTGCGAACCGGCTCGAGGAT---GCTCCGTTTCTGCGCGTCA 72278

QY 581 TCGTCCGCGCTCCATGCTGACCTCGATCGAAACCGGCGAGGAGCGGATCGGAAAGTTCG 640

Db 72279 ACCAGATGGCGTACCGCGGCTCTGTGCGAGGACGTGATGTCGGGCGGTGCGCATCC 72338

QY 641 CGGCGCGGCGGAGGCGCGGAGGCTCGGAGGCGGCTCGGAGGCGGCTCGGAGGCGGCT 697

Db 72339 GTGCGGCGGACAACTGATCCCGCTCTACTCGACGCGCAACCGCGACCCCGAGGTGT 72395

RESULT 11

US-09-103-840A-2

; Sequence 2, Application US/09103840A

; Patent No. 6294328

; GENERAL INFORMATION:

; APPLICANT: FLEISCHMAN, Robert D.

; APPLICANT: WHITE, Owen R.

; APPLICANT: FRASER, Claire M.

; APPLICANT: VENTER, John C.

; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

; TITLE OF INVENTION: TUBERCULOSIS

; FILE REFERENCE: 24366-20007.00

; CURRENT APPLICATION NUMBER: US/09/103,840A

; CURRENT FILING DATE: 1998-06-24

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 2

; LENGTH: 4403765

; TYPE: DNA

; ORGANISM: Mycobacterium tuberculosis

; FEATURE:

; OTHER INFORMATION: CDC 1551

; OTHER INFORMATION: "n" bases at various positions throughout the sequence

; OTHER INFORMATION: represent a, t, c or g

US-09-103-840A-2

Query Match 9.0%; Score 62.8; DB 3; Length 4403765;

[illegible]

```

RESULT 12
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

	Query Match	9.0%;	Score 62.8;	DB 3;	Length 4411529;
	Best Local Similarity	49.2;	Pred. No. 0.00063;		
	Matches 194;	Conservative 0;	Mismatches 197;	Indels 3;	Gaps 1;
Qy	124	GAGATCCTCTCGCGGGGCTCGGTCAACAAGGGCGCATGTACTTCCACTTTCGCTTCCAAG	183		
Db	3652184	GGGATTCGCGGATCGTCATGCTCACCGGACACACCGGAGATCGCCAGGTTGTCCG	3652243		
Qy	184	GAGAGCTGGCCCGCGCGTCTGGCCGAGCAGACCTGCACTGGCGGTGCGGAATCC	243		
Db	3652244	GACAACTGGGATGTATGATGTGCGCGCCGAGTCAATCCGAGGACAGCTCGCGCG	3652303		
Qy	244	GGCTCCAAAGGGCAGGAAGTGTGTAGACTTACCATGCTGCTCGCCACGGCATCTGCAC	303		
Db	3652304	GTGCGCAGTGTCCAGGACGACGGCTACGTCTCGGGATGTCGCGCAGCGCATCAACGAC	3652363		
Qy	304	GATCCGATCTCTCGGGGGGGGGC---ACCGCGTCTGCACCTGGACCAAGGGGGCGGTGGACTTC	360		

Db	3652364	GCCCCGGCGTGGCGCGCGCCGATATCGGGATCGCCATGGGCTTCGCCGAACCGACGCTC	3652423
Qy	361	TCCGAGCCAAACCCGTTTCGCGCAGTGGGGCGACATCTGCGCCACAGTCTCTGGCGGAGGCA	420
Db	3652424	GCGTTCGAGACCGCGCATGTCGCGCTGGCCAAACGACGACTGCAACGCCTGCTCGACGTT	3652483
Qy	421	CAGGAACGGGGGAGGTGTTCTCCGCACGTGAACCCGAAAAAGACCGCGCACTTCATCGTC	480
Db	3652484	GGGGACCTGGCGAGCGGGCAGTGGATGTAATCCGGCAGAACTACGGCATGTCCATCGCC	3652543
Qy	481	GGCTGCTTACCGGGGCTCCAGGCGGTCTCCCGG	514
Db	3652544	GTCAACCGGCGCGGGTGTGTGATCGGCGCGGGCG	3652577

RESULT 13
US-09-902-540-4889
; Sequence 4889, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 4889
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-4889

Query Match	8.9%	Score 62	DB 4	Length 630
Best Local Similarity	48.3%	Pred. No. 0.00026		
Matches	204	Conservative 0	Mismatches 215	Indels 3
Gaps	1			
QY	34	CAGGAACGGGCCCTCGCACGCGCAGGCGATCGTGCGGCAGCGCGCTCGGTCTTCGAC	93	
DB	37	CAGGAGCGCAGCCGCTCAGCGCCACGCGTCTGATGGAGCGGCCATTGGCGCGCTCTCC	96	
QY	94	GAGTACGGGTTTCAGGCGCGCCACAGTGGCAGAGATCCTCTCGGGGGCTCGGTACCAAG	153	
DB	97	GAGCTGGGGTGGCGGGCGCGACCATGCTGTTGATTCGAGCGCGCGGTGTGTCCCGG	156	
QY	154	GGCGCGATGTACTTCACATTCCTGGTTCACAGGAAGAGCTGGCCGCGCGCGTGTGGCCGAG	213	
DB	157	GGCGCCTGTTCAGCACCACTTCCCACCGCGCCGACTGTGTGGCCGCGCGGTGAGTAC	216	
QY	214	CAGACCTGACGTGGCGGTTCGCGAATCCGGCTCCAGCGCAGGAACTGTGTAGA---C	270	
DB	217	GTGGGGCACCAGCAGGTGGAGGGCGTGTCTCCGAAGGCCGCCAGTTGCCGCGCGATGCT	276	
QY	271	CTACCATGTGTGTGGCCACGCGATGCTGCACGATCCGATCTCTCGCGGGCGGACGCGG	330	
DB	277	CGCGCACGGAGGCCATCTCCACATGTGGCGGGCATCTACCTCAGCCCTGTCTTCACC	336	
QY	331	CTCGCATGGAACAGGGGGCGGTGGACTTCTCCGACGCCAACCCGTTTCGCGGAGTGGGGC	390	
DB	337	CGCGCGTGCAGTCTGTGGTGGCGCGCGTGGCGGACGCGGAGTGGCGCGCAGTTGGCC	396	
QY	391	GACATCTGCCCCAGCTCTCTGGCGGAGGACACAGAAACGGGGGAGGTGTCTTCGCGACGTG	450	
DB	397	CCGCTGGAGCCCGCGTGGGGCGGGAGTTCCACCGGCTGACGGTGCAGTTGCTCTCGGCGTG	456	
QY	451	AA	452	
DB	457	GA	458	

RESULT 14
US-09-902-540-1241
; Sequence 1241, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1241
; LENGTH: 28493
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(28493)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1241

Query Match 8.9%; Score 62; DB 4; Length 28493;
Best Local Similarity 48.3%; Pred. No. 0.00045; Mismatches 215; Indels 3; Gaps 1;
Matches 204; Conservative 0;
QY 34 CAGGAACGGCGGCTCCGACGCGGAGGCGATCGTCGGGGAGCGCCCTCGGTCTTCGAC 93
Db 6110 CAGGAGCGCAGCGCGTACGCGCAGCGTCTGATGGAGGCGGCCATTGGCGGCTCTCC 6169
QY 94 GAGTACGGTTGAGCCGCCACAGTGGCAGAGATCTCTTCGCGGCGCTTCGTTACCAAG 153
Db 6170 GAGCTGGGGTGGCGGGCGCGACCATGACTGTGATTCGCGAGCGCGGGTGTGCCCG 6229
QY 154 GCGCGATGCTACTTCCACTTCCTCCAGAGAGAGTGGCGCGCGGCGTGTGGCGCGAG 213
Db 6230 GCGCGCTGTGAGCACCATTCCTCCACCGCGCGGACCTGTGGCGCGCGCGGCGGATC 6289
QY 214 CAGACCTGACGTGCGGCTGCGGAATCCGGCTCCAGGCGCAGGAATGTTAGTAA--C 270
Db 6290 GTGGGACACAGCAGGTGAGCGGTGCTCCGCAAGGCGCCAGTTGCGGCGCGATGCT 6349
QY 271 CTCACCATGCTGTGCGCCACCGCATCTGTCAGAGATCCGATCTGCGGGCGGGCAGCGG 330
Db 6350 CGCCGACAGGAGGCCATCTCCACATGCTGGCGCGGCATCTACCTCAGCCCTGTCTTACC 6409
QY 331 CTCGCACTGGACAGGGGCGGTGGACTTCTCCGACGCCAACCCGTTGCGCGAGTGGGCG 390
Db 6410 GCGGGGTCAGCTCTGGGTGGCGCGGTGGCGAGCGGAGCTGCGCGCGAGTTGGCC 6469
QY 391 GACATCTGCGCCAGCTCTGCGGAGGACAGGAACGCGGGGAGGTGTTCCGACAGCTG 450
Db 6470 CCGTGGAGCGCGGTGGGGCGGAGTTCCACCGGCTGAOGTGCAGTTGCTCGCGGTG 6529
QY 451 AA 452
Db 6530 GA 6531

RESULT 15
US-09-773-816-1
; Sequence 1, Application US/09773816
; Patent No. 6340774
; GENERAL INFORMATION:
; APPLICANT: Stanford University
; APPLICANT: Khosla, Chaitan
; TITLE OF INVENTION: NON-STEROIDAL ESTROGEN-RECEPTOR
; TITLE OF INVENTION: ANTAGONISTS

FILE REFERENCE: 28600-20210.00
; CURRENT APPLICATION NUMBER: US/09/773,816
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/243,458
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: US 60/179,305
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 23673
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(23623)
; OTHER INFORMATION: n = A,T,C or G
US-09-773-816-1
Query Match 8.6%; Score 60.2; DB 3; Length 23673;
Best Local Similarity 45.0%; Pred. No. 0.0011; Mismatches 373; Indels 9; Gaps 2;
Matches 313; Conservative 0;
QY 14 ACGAACGGGTGGCAGTGGCAGCAGGAACGGGGCGCTCCGACGCGGCGAGTCTGTCGGG 73
Db 15958 ACACGCGGTGGCGCTGTACGTGCACGCGCTGGGCGTCAACCGGTGTGCGACGGTCCGAGT 16017
QY 74 CAGCGCGCTCGGTCTTCAGCAGATACGGGTTCGAGGCGCCACACATGTCGAGAGATCTCTT 133
Db 16018 TCGACGCTCCGCGCGACCGACGAGGACATGACCGCGGAGAGTGCAGGCGCAGCATCC 16077
QY 134 CGCGGGCTCGGTTCACCAAGCGCGGTGTACTTCCACTTCGCTTCCCAAGGAAGAGCTGG 193
Db 16078 GACGGGTGAGCGGACGAGCGTGCAGATCAGGAGATGTCTTCGGGCAACCGCTGATCG 16137
QY 194 CCGCGCGGTGCTGCGCGAGCAGACCTTCGACGTGGCGGTGCGCGAATCCGCTCCAAAG 253
Db 16138 ACACGCGCGGCGAGCGGTCCGCTACCGCTCGGCGGGTGTGCTGGCGCGGCGAGCGG 16197
QY 254 CGCAGGAAGTGTAGACCTCACCATGTGTGTCGCCACCGCATGTCGACGATCCGATCC 313
Db 16198 CGCAGTGTACGCGCGGTGCGCGCGCGAGGCGCTCAACGTGCGGCTGTGTCGACGCGGCA 16257
QY 314 TCGCGCGCGGACGCGGTGCTGACCTGACACGAGGGCGGTGAGCTTCTCCGACGCAACC 373
Db 16258 ACCTGCTGTGAGAGTGTGCGCGGCGAGGTACAGCGTGGCGCGCACTATCTCTCGACA 16317
QY 374 CGTTCGCGAGTGGGGCGACATCTGCGCCAGTCTTCGCGGAGGCGACAGGAACGCGGGG 433
Db 16318 CTTACGTTTCGAGAGCGGTACCGGTGCGCGCGCGCTTTTCAGAAACACCGCGCGGAGA 16377
QY 434 AGGTGCTTCCGACGTGAACCGGAAGAGACCG-----GGACTTCACTCGTGGCTGCT 487
Db 16378 TCGCGCTCATGACACCGCGCGCGAGACACCGCGCTGCGCGAGATGTTTCAGGAATGTC 16437
QY 488 TCACCGGCTCCAGCGGTCTCCCGGGTCACTTCGACCGCGAGACCTTCGCGCACCGGAG 547
Db 16438 TCGACATCAGCAGGTGTCACCGCAGATCGCGGACATGATGCCGGAATGACGTCGCT 16497
QY 548 TCTCGGTGATGTGAAACCAAGTGTGCCCCAGCA---TCGTGCGGCGTCTCATGTCGACT 604
Db 16498 ACGCCGTCGAGGGGAGCATCCGCTCTCTCGGACGCTGTCGCGGAGCGCAAGCTCACCG 16557
QY 605 GGATCGAAACCGGAGGAGGCGGATCGGGAAGTTCGCGGCGCGCGCGCGCGCGCGAGG 664
Db 16558 GCGCAAGCGCGCGCACTGTGGGCGGACGCGGCGGCGGTGCTGTCGACTTCGCGCGAC 16617
QY 665 CTGCGGAGGCTCCGAGCGCGCTCCCGACGAGTAG 699
Db 16618 GCGGCGAGTGGCGGACGCGGTTCGCGCGGTGCTG 16652

Search completed: March 6, 2005, 22:16:10

Mon Mar 7 10:28:17 2005

us-10-049-710a-1.rni

Page 9

Job time : 242 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2005, 19:57:05 ; Search time 630 Seconds
(without alignments)
6583.744 Million cell updates/sec

Title: US-10-049-710A-1
Perfect score: 699
Sequence: 1 atggcagtcgcacagaacg.....aggccgcctccagcagtag 699

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5401638 seqs, 2966923429 residues

Total number of hits satisfying chosen parameters: 10803276

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	699	100.0	699	15	US-10-049-710A-1
2	229	32.8	4346	15	US-10-017-471A-19
3	157.4	22.5	702	15	US-10-156-761-3692
4	157.4	22.5	9025608	15	US-10-156-761-1
5	112.4	16.1	9025608	15	US-10-156-761-1
6	112.2	16.1	654	15	US-10-156-761-1
7	93.4	13.4	654	15	US-10-156-761-3689
8	85.8	12.3	9521	17	US-10-168-663-18
9	85.8	12.3	9521	17	US-10-168-663-18
10	78.6	11.2	851	15	US-10-156-761-3690
11	78.2	11.2	4257	9	US-09-825-288A-1

12	69.6	10.0	18438	15	US-10-156-761-2886	Sequence 2886, Ap
13	69.6	10.0	125746	15	US-10-156-761-15102	Sequence 15102, A
14	68.4	9.8	3173	18	US-10-437-963-77080	Sequence 77080, A
15	68.2	9.8	1224	17	US-10-282-1302A-25841	Sequence 25841, A
16	68.2	9.8	32329	17	US-10-374-903A-1	Sequence 1, Appli
17	67.2	9.6	1041	9	US-09-815-242-7885	Sequence 7885, Ap
18	67.2	9.6	1041	17	US-10-282-122A-30427	Sequence 30427, A
19	66.4	9.5	603	15	US-10-156-761-4514	Sequence 4514, Ap
20	65	9.3	657	15	US-10-156-761-5897	Sequence 5897, Ap
21	65	9.3	77536	10	US-09-940-316B-1	Sequence 1, Appli
22	64	9.2	9521	17	US-10-168-663-18	Sequence 18, Appl
23	64	9.2	9521	17	US-10-168-663-19	Sequence 19, Appl
24	63.4	9.1	5862	16	US-10-132-134-15	Sequence 15, Appl
25	63.4	9.1	52101	16	US-10-132-134-1	Sequence 1, Appli
26	63.2	9.0	681	15	US-10-156-761-6081	Sequence 6081, Ap
27	63	9.0	591	15	US-10-156-761-4766	Sequence 4766, Ap
28	62.8	9.0	618	15	US-10-156-761-2255	Sequence 2255, Ap
29	62.6	9.0	582	17	US-10-282-122A-11537	Sequence 11537, A
30	62.6	9.0	633	15	US-10-156-761-558	Sequence 558, App
31	62.4	8.9	819	17	US-10-282-122A-11538	Sequence 11538, A
32	62.4	8.9	109519	11	US-09-758-759-122	Sequence 122, Appl
33	62.4	8.8	651	15	US-10-156-761-744	Sequence 744, App
34	61.4	8.7	3516	18	US-10-437-963-74025	Sequence 74025, A
35	60.8	8.7	5877	14	US-10-152-886-94	Sequence 94, Appl
36	60.8	8.7	381	18	US-10-021-323-17015	Sequence 17015, A
37	60.6	8.7	1395	15	US-10-156-761-2088	Sequence 2088, Ap
38	60.2	8.6	1836	15	US-10-156-761-2659	Sequence 2659, Ap
39	60.2	8.6	27541	18	US-10-203-295-2	Sequence 2, Appli
40	60.2	8.6	125401	18	US-10-203-295-35	Sequence 35, Appl
41	60.2	8.6	1893	15	US-10-156-761-5675	Sequence 5675, Ap
42	60	8.6	27541	18	US-10-203-295-2	Sequence 2, Appli
43	60	8.6	125401	18	US-10-203-295-35	Sequence 35, Appl
44	59.8	8.6	573	15	US-10-156-761-4685	Sequence 4685, Ap
45						

ALIGNMENTS

RESULT 1

US-10-049-710A-1
; Sequence 1, Application US/10049710A
; Publication No. US20030126648A1
; GENERAL INFORMATION:
; APPLICANT: Shimmyo, Atsuhiko
; APPLICANT: Kato, Kou
; APPLICANT: Yamada, Yasuhiro
; APPLICANT: Nihira, Takuya
; APPLICANT: Shindo, Takuya
; TITLE OF INVENTION: METHOD FOR INDUCTION OF GENE EXPRESSION IN PLANT AND PLANT TREAT
; FILE REFERENCE: 5405/18
; CURRENT APPLICATION NUMBER: US/10/049,710A
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: PCT/JP01/05096
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: JP 2000-180466
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 699
; TYPE: DNA
; ORGANISM: Streptomyces virginiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(699)
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Okamoto, S., Nakamura, K., Nihira, T. and Yamada, Y.
; TITLE: Virginiae butanolide binding protein from Streptomyces virginiae.
; TITLE: Evidence that VbrA is not the virginiae butanolide binding protein and re-
; TITLE: Identification of the true binding protein

JOURNAL: Journal of Biological Chemistry

VOLUME: 270

ISSUE: 20

PAGES: 12319-12326

DATE: 1995-05-19

DATABASE ACCESSION NUMBER: D32251

DATABASE ENTRY DATE: 1994-07-19

RELEVANT RESIDUES: (1)..(699)

PUBLICATION INFORMATION:

AUTHORS: Okamoto, S., Nakamura, K., Nihira, T. and Yamada, Y.

TITLE: Virginiae butanolide binding protein from Streptomyces virginiae.

TITLE: Evidence that VbrA is not the virginiae butanolide binding protein and re-

TITLE: Identification of the true binding protein

JOURNAL: Journal of Biological Chemistry

VOLUME: 270

ISSUE: 20

PAGES: 12319-12326

DATE: 1995-05-19

DATABASE ACCESSION NUMBER: D32251

DATABASE ENTRY DATE: 1994-07-19

US-10-049-710A-1

Query Match 100.0%; Score 699; DB 15; Length 699;

Best Local Similarity 100.0%; Pred. No. 6e-172;

Matches 699; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGCAGTCCGACACGAAACGGGTGGCAGTGGCGACAGGAAACGGCGCTCGCACGGCGGACG 60
DB 1 ATGGCAGTCCGACACGAAACGGGTGGCAGTGGCGACAGGAAACGGCGCTCGCACGGCGGACG 60
QY 61 GCGATCGTGGGCGACGCCCTCGGTCTTTCGACGAGTACGGGTTCGAGGCCGCCACAGTG 120
DB 61 GCGATCGTGGGCGACGCCCTCGGTCTTTCGACGAGTACGGGTTCGAGGCCGCCACAGTG 120
QY 121 GCAGAGATCCTCTCGGGGCGCTCGGTACCAAGAGGGCGGATGACTTCCACTTCGGTTCC 180
DB 121 GCAGAGATCCTCTCGGGGCGCTCGGTACCAAGAGGGCGGATGACTTCCACTTCGGTTCC 180
QY 181 AAGGAAGAGTGGCGCGCGCGTGGTGGCGAGCAGACCTCGACGTGGCGGTGGCGGAA 240
DB 181 AAGGAAGAGTGGCGCGCGCGTGGTGGCGAGCAGACCTCGACGTGGCGGTGGCGGAA 240
QY 241 TCCGGTCCAAAGCGCGAGAACTGGTACCTTACCATGCTGGTGGTGGTGGTGGTGGTGGTGG 300
DB 241 TCCGGTCCAAAGCGCGAGAACTGGTACCTTACCATGCTGGTGGTGGTGGTGGTGGTGGTGG 300
QY 301 CACGATCCGATCTCGGGGCGGCGACGCGGCTCGCACTGGACCGAGCGGCGGTGGACTTC 360
DB 301 CACGATCCGATCTCGGGGCGGCGACGCGGCTCGCACTGGACCGAGCGGCGGTGGACTTC 360
QY 361 TCCGAGCGCAACCCGTTCCGGGAGTGGGCGACATCTGGCGCCAGCTCTTGGCGGAGGCA 420
DB 361 TCCGAGCGCAACCCGTTCCGGGAGTGGGCGACATCTGGCGCCAGCTCTTGGCGGAGGCA 420
QY 421 CAGGAACCGGGGAGGTGTTCCGCACTGTAACCCGAAAGAACCGGCGACTTCATCGTC 480
DB 421 CAGGAACCGGGGAGGTGTTCCGCACTGTAACCCGAAAGAACCGGCGACTTCATCGTC 480
QY 481 GGTGCTTCCAGGGGCTCAGGGGCTCCGGGTCCCTCGGACCGCGGAGCGGCGGCGG 540
DB 481 GGTGCTTCCAGGGGCTCAGGGGCTCCGGGTCCCTCGGACCGCGGAGCGGCGGCGGCGG 540
QY 541 CACCGGATCTCGGTGATGTGGAAACACAGCTGCTGCCAGCATCGTGGCGGCGTCCATGCTG 600
DB 541 CACCGGATCTCGGTGATGTGGAAACACAGCTGCTGCCAGCATCGTGGCGGCGTCCATGCTG 600
QY 601 ACTTGATCGAAACCGCGGAGAGCGGATCGGAAGGTTCGGGCGGCGGCGGCGGCGGCGG 660
DB 601 ACTTGATCGAAACCGCGGAGAGCGGATCGGAAGGTTCGGGCGGCGGCGGCGGCGGCGGCGG 660
QY 661 GAGGCTCGGAGGCTCCGAGGCGGCTCCGACGAGTAG 699
DB 661 GAGGCTCGGAGGCTCCGAGGCGGCTCCGACGAGTAG 699
```

RESULT 2

US-10-017-471A-19

Sequence 19, Application US/10017471A

Publication No. US2003012464A1

GENERAL INFORMATION:

APPLICANT: Takano, Eriko

APPLICANT: Bibb, Mervyn

TITLE OF INVENTION: Antibiotic Production

FILE REFERENCE: 0380-P02329US1

CURRENT APPLICATION NUMBER: US/10/017,471A

CURRENT FILING DATE: 2001-10-23

PRIOR APPLICATION NUMBER: US 60/242,561

PRIOR FILING DATE: 2000-10-23

NUMBER OF SEQ ID NOS: 19

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 19

LENGTH: 4346

TYPE: DNA

ORGANISM: Streptomyces coelicolor

US-10-017-471A-19

Query Match 32.8%; Score 229; DB 15; Length 4346;

Best Local Similarity 60.8%; Pred. No. 5.2e-50;

Matches 373; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

```
QY 34 CAGGAACGGGCGTCCGCGACGGCGGCGGATCGTGGCGGCGAGCCCTCGGTCTTCGAC 93
DB 2270 CAGGAACGGGCGATCCGCGCGGCGGATCGTGGCGGCGGCGGCGGCGGCGGCGGCGG 2329
QY 94 CAGTACGGGTTTCGAGCGCGGCCACAGTGGCAGAGATCTCTCGGGGCGCTCGGTACCAAG 153
DB 2330 AAGCAGGCTACCAAGCTGCCAGATCAGGAGATCTCAAGTGGCGGCGGCGGCGGCGG 2389
QY 154 GCGCGATGTACTTCCACTTCGCTTCCAGAGAGAGTGGCGGCGGCGGCGGCGGCGGCGG 213
DB 2390 GCGAGCCCTTACTTCCACTTCAGTCCAGGAGAACTGGCGCTGGGCGGTCTTCGACGCC 2449
QY 214 CAGACCTTGCACGTGGCGGTGGCGAATCGCGCTCCAAGGCGCAGAACTGGTAGACCTC 273
DB 2450 CAGGAACCAACACAGCGGCTTCGAGGCAACCCCTCCGGCTGCAAGAACTCATCGACATG 2509
QY 274 ACCATGTGTGTCGCCACGGCATGTGCAGATCCGATCTCTCGGGGCGGCGACGCGGCTC 333
DB 2510 GGCATGTTGTTCTGTACCGCTTGGCGACGAACTGGTGGCGGCGGCGGCGGCGGCGGCTC 2569
QY 334 GCACGTGGACAGGGGCGGTGGACTTCTCCGAGCGCAACCCGTTCCGGGAGTGGGCGGAC 393
DB 2570 TCCATGGACAGCGGCGGCGGCTTCGATCCGAGGACCTTCGCTCGTGGCGACGAG 2629
QY 394 ATCTCGCGCCAGCTCTCGCGGAGGCGACAGGAAACGGGGGAGGTGCTTCGCGACGTTGAAC 453
DB 2630 ACATCTCTGAAGTCTGTAACCGAGCGCAAGAGACGAGTGTGCTGCCCATGTGGTC 2689
QY 454 CCGAAAAAGACCGCGACTTCATCTCGGTGCTCTTACCGGCTCCAGCGGCTCCCGG 513
DB 2690 ACCACCGACTCGCGCGATCTCTACGTGGCGACGTTCCGCGGATACAGGTCTGTGTCGCG 2749
QY 514 GTCACTCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 573
DB 2750 ACGGTGACGACATACAGGACCTCGAAACCGCTACCGGCTGCTGCGAAGACATCTCTG 2809
QY 574 CCGAGCATCGTCCGGCGGCTCCATGCTGACTCGAATCGAAGCGGCGGAGGCGGAGTCGGG 633
DB 2810 CCGCGCATCGGCTTCCCTCCGCTGGCGGCGGCTGCTCTCCGAGGAGCGGCGGAGCA 2869
QY 634 AAGGTGCGGCGG 646
DB 2870 CGCCTCGGCGCG 2882
```

RESULT 3

US-10-156-761-3692
; Sequence 3692, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 3692
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(702)
US-10-156-761-3692

Query Match 22.5%; Score 157.4; DB 15; Length 702;
Best Local Similarity 55.8%; Pred. No. 2.2e-31;
Matches 371; Conservative 0; Mismatches 276; Indels 18; Gaps 3;
QY 28 GTGCGACAGAAACGGGCGCTCCGACGCGGCGAGCATCGTGGGCGAGCGCGCTCGGTC 87
DB 4 GCGCGGACGAGGAGCGCCATTCCGACGCGGAGCATCTGTGTCGCGGCGCGAGGTG 63
QY 88 TTCGACGAGTACGGGTTGAGGCGGCGACAGTGGGAGAGATCTCTCGGCGGCGCTCGGTC 147
DB 64 TTCGACGAGTGGGATACGAGGCGGCAACCATCTCCGAGCTGTGAAGCGCTCGGGGTC 123
QY 148 ACCAAGGGCGGATGACTTCCACTTCGTTCCAAAGAGAGAGTGGCCGCGCGGTGCTG 207
DB 124 ACCAAGGGGCGCTCTACTTCCACTTCAGTCGAAGCAGAGTGGCCCGCGGCTGCTG 183
QY 208 GCGGACGAGACCTGACAGCTGGCG---GTGCGCGAATCCGGCTCCAAAGCGCAGGAAC 264
DB 184 GCGGACGAGTCCGCTCCCTTCGCGGTCGCCGAGCAGGAGCTGAAGCTCCAGCAGTCG 243
QY 265 GTAGACCTCACCAGTGTGTCGCGGCGAGGATGCTGCA-----CGATCCGATCCTG 315
DB 244 CTGGACGAGGCGTGTGCTCGGCCATCTGCTCAGGGAAGGCACCGCGCGATCCGATCGTC 303
QY 316 CGGCGGCGACGCGGCTCCGACTGGACACAGGGGGC-----GTGGAATCTCCGAGCC 369
DB 304 CAGGCGAGTGTGGGCTGACCGTGGACAGGGCTCCCGAGGAGCCATCTCAACCGCGCG 363
QY 370 AACCCCTTCGCGAGTGGGCGGACATCTGCGGCCAGCTCTGCGGAGGACAGGAACCG 429
DB 364 GTCCCGATCGAGGCTGGACCGAGCACAGCAGTCCCTCTTGAAGAGGCGAGGCGCAAG 423
QY 430 GGGGAGGTGTTCCGCGAGTGAACCCGAAAGACCGGCGACTTCATGTCGCGCTGCTTC 489
DB 424 GCGGAGATCTGCGCCACGCGGATGTGAAGCGCTCGCCAAAGCTGTGTCGCGCGGCTTC 483
QY 490 ACCGGGCTCCAGGCGTCTCCCGGGTCACTCCGACGCGCAGGACCTCGGCCACCGGATC 549
DB 484 ACCGGGCTGAGGCTCTCTGAGGATCATGACCGGGCGCGGAGCTGGCGGAGCGGGTG 543
QY 550 TCGGTGATGTGAACACAGCTGTGCGCAGCATCGTCCGCGCGTCCACTGTGACTGAGCTG 609
DB 544 GCGGACCTCTACCGCATCTGATCGCGTCTTCGCCATGCGGGGATCTGTGTCGCGCTG 603

QY 610 GAAACCGGCGAGGAGCGGATCGGAAGTTCGCGGCGGCGCGAGCGCCCGCGAGGCTGCG 669
DB 604 GACTTCTCCCGGAGCGGGGCTCGCGGCTGTACGAAGCCCGCATGAAGCAGCGGAGTGC 663
QY 670 GAGGC 674
DB 664 GCGGC 668
RESULT 4
US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1
Query Match 22.5%; Score 157.4; DB 15; Length 9025608;
Best Local Similarity 55.8%; Pred. No. 1.2e-31;
Matches 371; Conservative 0; Mismatches 276; Indels 18; Gaps 3;
QY 28 GTGCGACAGAAACGGGCGCTCCGACGCGGCGAGCATCGTGGGCGAGCGCGCTCGGTC 87
DB 4585273 GCGCGGACGAGGAGCGCCATTCCGACGCGGCGAGCATCTGTGTCGCGCGCGAGGTG 4585214
QY 88 TTCGACGAGTACGGGTTGAGGCGGCGCACAGTGGGAGAGATCTCTCGCGGGGCTCGGTC 147
DB 4585213 TTCGACGAGTGGGATACGAGGCGGCAACCATCTCCGACGCTGTGAAGCGCTCGGGGTC 4585154
QY 148 ACCAAGGGCGGATGACTTCCACTTCGTTCCAAAGAGAGAGTGGCCGCGCGGTGCTG 207
DB 4585153 ACCAAGGGGCGCTCTACTTCCACTTCAGTCGAAGCAGAGTGGCCCGAGGCGGCTG 4585094
QY 208 GCGGACGAGACCTGTCAGCTGGCG---GTGCGCGAATCCGGCTCCAAAGCGCAGGAAC 264
DB 4585093 GCGGACGAGTCCGCTCCCTTCGCGGCTCCCGAGCAGGAGTGAAGCTCCAGCAGTCG 4585034
QY 265 GTAGACCTCACCAGTGTGTCGCGGCGAGGATGCTGCA-----CGATCCGATCCTG 315
DB 4585033 CTGGACGAGGCGCTGCTGCTCGCCATCTGCTCAGGGAAGGCACCGCGCATCCGATCGTC 4584974
QY 316 CGGCGGCGCACGCGGCTCGCACTGGACACAGGGGGC-----GTGGAATCTTCGACGCC 369
DB 4584973 CAGGCGAGTGTGCGGCTGACCTGGACGAGGCTGCCCGAGGAGCCATCTCAACCGCGG 4584914
QY 370 AACCCGTTTCGCGAGTGGGCGGACATCTGCGGCCAGCTCTCGCGGAGGACACAGGAACCG 429
DB 4584913 GTCCGATGACGAGGCTGGACCGGAGCACAGCAGTCCCTCTTGAAGAGGCGGCGCAAG 4584854
QY 430 GGGGAGGTGCTTTCGCGCATGTAACCCGAAAGACCGGCGACTTTCATGTCGCGCTGCTTC 489

Db 4584853 GCGGAGATCTCTCCACCGCCGATGTGGAAGCGCTCGCAAGCTGTGTGCGGCGGTTTC 4584794
 QY 490 ACCGGGCTCCAGCGGCTCTCCGGGTACCTCGACCGCCAGGACTCGGCCACCGGATC 549
 Db 4584793 ACCGGGCTCAGGTCTCTCGAGGATCATGACCGGCGCGGAGCTCGCGGAGCGGGTG 4584734
 QY 550 TCGGTGATGTGGAACACAGTGTGCCAGCATCTGCGGGCTTCATGCTGACCTGGATC 609
 Db 4584733 GCGGACCTTACCGCATCTGATGCGTCTTGCCATGCCGGGATCTGTGCGCCTG 4584674
 QY 610 GAAACCGGAGAGCGGATCGGGAAGGTTCGGCGCGCGCGAGCCCGAGGCTGG 669
 Db 4584673 GACTTCTCCCGAGCGGGCTCGCGGTGTACGAAGCCGCATGAAGCAGCGGAGTGC 4584614
 QY 670 GAGGC 674
 Db 4584613 GCGGC 4584609

RESULT 5

US-10-156-761-1
 ; Sequence 1, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; PRIOR FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 1
 ; LENGTH: 9025608
 ; TYPE: DNA
 ; ORGANISM: Streptomyces avermitilis
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (4187715)
 ; OTHER INFORMATION: a, t, c, g, other or unknown
 US-10-156-761-1

Query Match 16.1%; Score 112.4; DB 15; Length 9025608;
 Best Local Similarity 50.1%; Pred. No. 5.9e-20; Mismatches 326; Indels 15; Gaps 2;
 Matches 343; Conservative 0;
 QY 4 GCAGTGCACACCAACAGGGTGGCAGTGCACAGGAACAGGGCGCTCGCGACCGCGGAGCGG 63
 Db 4581408 GCCTTCGGGAGAGTCAACGGTGCACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4581467
 QY 64 ATCGTGGGGCAGCGGCTCGCTTCGACGAGTACGGGTTGAGGCGCGCACAGTGGCA 123
 Db 4581468 CTCATCGCTTCGCGCGCCACCGCTTCGAACGCGAGGCTTACACACAGGCGAGGCTGGCC 4581527
 QY 124 GAGATCTCTCGCGGGCTCGCTCACCAGGGCGGATGTACTTCCACTTCGCTTCCAAG 183
 Db 4581528 GACATGAGCCCTGCGCGGCTGTGAGCCCGGCGACTGCACTTCACTTCGAGAGCAAG 4581587
 QY 184 GAGAGGTGCCCGCGCGCTGCTGGCGGAGACACCTGCACTGCGGGTTCGCGGAATCC 243
 Db 4581588 GCAGAGTGCACAGGGCGGTGGAGGGCGGCGGGGTGAGCCTCGCGCGCGGCGCTGG 4581647
 QY 244 GGTCCAGGCGCAGGAACTGGTAGACCT-----CACCATGCTGGTGGCCAC 291

Db 4581648 CTGCCCCAGCGCGCGGACGAACGCGCTGCAACGGCTGACAAACACATCTGCACGCCCTG 4581707
 QY 292 GGCATGCTGCAGATCCGATCTCTGCGGCGCGGACGCGGCTCGCACTCGACACGAGGGGGG 351
 Db 4581708 GCGGAGCAGTTCGCGGGGAGAGTGTGTCGCCCGCGGGGCTTCGGGCTGAACTGCCAATCG 4581767
 QY 352 GTGACTTCTTCGAGCCCAACCCGTTCCGCGAGTGGGCGGACATCTGCGGCCAGCTCTCTG 411
 Db 4581768 GCGGCGCGCGCGCTGAACTCTGCTCCGGAATGCGAGACCTGCTGTGAGCAGCTGTCTC 4581827
 QY 412 GCGGAGGACAGGAACGCGGGAGGTGCTTCGCGACGTGAACCCGAAAGAACCGCGGAC 471
 Db 4581828 GCGGAGGCGCGGAGGAGGGGCTGCTGCCCGGCGCTTCGTCGCGCGGACACGCTCAGC 4581887
 QY 472 TTCATGCTGCGTGTCTTACCGGGCTCCAGGCGGCTCTCCCGGCTCACCTCCGACCGCCAG 531
 Db 4581888 GCGGTGTGGCGCGACCAACCGGTTTCGAACTGCTCGCGCGCGGACCCCGAGTGGCT- 4581946
 QY 532 GACCTCGGCCACCGGATCTCGGTGATGTGGAACCAACGTCGTGTCGCCAGCATCTGTCCGGCG 591
 Db 4581947 --CTCGGCCAGTCTGCTGCGCGCTTCTGCGGGTACTGCTGCCGCGCGCGCGCGG 4582004
 QY 592 TCCATGCTGACTGATCGAAACCGCGGAGGAGCGGATCGGGAAGTCTCGCGCGCGGCC 651
 Db 4582005 GCGGCCCTGACCGCGGTGACCCCGGACGCGGAGCTGCCACCGCGCGGAGACCGGACC 4582064
 QY 652 GAGGCGCGGAGCTGCGGAGGCC 675
 Db 4582065 CCGGCGACCAACCGCGCGGATGAGCC 4582088

RESULT 6

US-10-156-761-3689
 ; Sequence 3689, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 3689
 ; LENGTH: 654
 ; TYPE: DNA
 ; ORGANISM: Streptomyces avermitilis
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(654)
 US-10-156-761-3689

Query Match 16.1%; Score 112.2; DB 15; Length 654;
 Best Local Similarity 50.9%; Pred. No. 1.2e-19; Mismatches 303; Indels 15; Gaps 2;
 Matches 329; Conservative 0;
 QY 33 ACAGGAACGGCGCTCGCGACCGCGGAGGAGTTCGTGCGGCGAGCGCGCTCGGTCTTCGA 92
 Db 9 ACAGGAACGGCGCGCGCGGACCGCGCACCGCTTCATCCGCTCGCGCGCCACCGCTTCGA 68
 QY 93 CGAGTAGCGGTTTCGAGGCGCGCACAGTGGCAGAGATCTCTTCGCGGGGCTTCGGTTCACCAA 152
 Db 69 ACGGCAAGGCTACACACAGGCGAGGCTGCGCGGACATCAGCGCTGCGCGGCTGTGAGCCC 128

90 CGACGAGTACGGGTTTCGAGGCGCGCCACAGTGCAGAGATCTCTCGCGGCGCTCGGTAC 149
 4932 CGCTCGCGCGCTACCGAGGGCCCTCGTCAAGGACGTTGCGGAGCGTGTGCGGATGAC 4991
 150 CAAGGGCGGATGATCTTCATTCGCTTCAAGGAGAGTGTGGCGCGGCGGTCTGCGC 209
 4992 CAAGGGCGGATGATCTTCATTCGCTTCAAGGAGAGTGTGGCGCGGCGGTCTGCGC 5051
 210 CGACGAGACCTCGACGTTGGCGG-----TGCCGGAATCCGGCTCCCAAGGC----- 254
 5052 CGAGCACTACGCGCGCTGCGCCGCGAGCGATGGAAGAGATCCGCATCCAGGGCTTCACACC 5111
 255 ---CGAGGAATCGGTAGACTCACCATGCTGTGTCGCCACGCGCATGCTGCAGATCCGAT 311
 5112 GCTGAGACGCTCGAGGATGCTCCATCGCGCGGCGAGGCTTCCGCGAGACCCCGT 5171
 312 CTTGCGGGCGGCGACGCGGCTCGCATGTGACACAGGGGGCGGTGAGACTTCTCCGAGCCAA 371
 5172 GATGACAGCGGTGCGCGGCTCGAGAGTGAAGCGCGCTTTCATCGACGCGGAGCTGCCCT 5231
 372 CCGGTTCCGCGAGTGGGGGACATCTGCGCCCGAGCTCTGCGGAGGACACAGGAACGGGG 431
 5232 GCCCTACGTGGAGTGAACCCACTGCTGAGGTGCGGTTGAGGAGCGCCCGTGAAGCGCG 5291
 432 GGAGGTGCTTCCGCGACGTGAACCCGAAAGACCGGCGGACTTTCATCGTGGCTGCTTCAC 491
 5292 CCAGTTGCGGGCGGCGTGCATCCCGCAGCAGCTGCGCGTTCCTGTTGCGCGCTTCTT 5351
 492 CGGGTCCAGGCGGTCTCCGGGTACCTCCAGCGCGGAGGACCTGCGGCGACCGGATCTC 551
 5352 CGGCATGACAGCAGCTCTCCGACAATCTGCACAGCGAGCGGACATCATGAGCGGTGGCA 5411
 552 GGTGATGTGGAACACGCTGCT 572
 5412 GGAGCTGCGGGAGCTGATGTT 5432

RESULT 9
 US-10-168-663-19/c
 ; Sequence 19, Application US/10168663
 ; Publication No. US2004008692A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Plant Biotechnology Limited
 ; APPLICANT: Chater, Keith F
 ; APPLICANT: Bruton, Celia J
 ; APPLICANT: O'Rourke, Sean J
 ; APPLICANT: Wietzorrek, Andreas W
 ; TITLE OF INVENTION: Methods and Materials Relating to Gene Expression
 ; FILE REFERENCE: 0380-P02909US0
 ; CURRENT APPLICATION NUMBER: US/10/168,663
 ; PRIOR FILING DATE: 2002-10-25
 ; PRIOR APPLICATION NUMBER: PCT/GB00/04972
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: GB 9930477.6
 ; PRIOR FILING DATE: 1999-12-23
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 19
 ; LENGTH: 9521
 ; TYPE: DNA
 ; ORGANISM: Streptomyces coelicolor
 US-10-168-663-19
 Query Match 12.3%; Score 85.8; DB 17; Length 9521;
 Best Local Similarity 50.1%; Pred. No. 7.3e-13;
 Matches 281; Conservative 0; Mismatches 262; Indels 18; Gaps 2;
 30 GCGACAGGAACGGGCGCTCCGCGAGGCGGATCGTGGCGGCGAGCCCGCTCGGTCTT 89
 4650 GCAGCAGGAGCGGTTCGATCAAGACCCCGGCCAGATCCTTGGAGCGCGCTCGGAGATCTT 4591
 90 CGACGAGTACGGGTTCGAGGCGCGCCACAGTGGCGAGATCTCTCGCGGCGCTCGGTAC 149

Db 4590 CGCTCGCGCGGCTTACCGAGGGGCGCTCGGTCAAGGAGCTTCCCGAGCGGTGTGCGGATGAC 4531
 Qy 150 CAAGGGCGGATGATCTTCATTCGCTTCAAGGAGAGTGTGGCGCGGCGGTCTGCGC 209
 Db 4530 CAAGGGCGGATGATCTTCATTCGCTTCCCGAGCAAGGAATCACTGGCCATCGCGGTGTTGA 4471
 Qy 210 CGAGCAGACCTCGACGTTGGCGG-----TGCCGGAATCCGGCTCCCAAGGC----- 254
 Db 4470 GGAGCACTACGCGCGCTGCGCCGCGAGCGATGGAAGAGATCCGCATCCAGGGCTTCACACC 4411
 Qy 255 ---CGAGGAATCGGTAGACTCACCATGCTGTGTCGCCACGCGCATGCTGCAGATCCGAT 311
 Db 4410 GCTGAGACGCTCGAGGAGATGCTCCATCGCGCGGCGAGGCTTCCGCGAGACCCCGT 4351
 Qy 312 CTTGCGGGCGGCGACGCGGCTCGCATGTGACACAGGGGGCGGTGGAATCTTCCGAGCCAA 371
 Db 4350 GATGACAGCGGTGCGCGGCTGCAAGTGAAGCGCGCTTTCATCGACGCGGAGCTGCCCT 4291
 Qy 372 CCGGTTCCGCGAGTGGGGGACATCTGCGCCCGAGCTCTGCGGAGGACACAGGAACGGGG 431
 Db 4290 GCCCTACGTGGAGTGAACCCACTGCTGAGGTGCGGTTGAGGAGCGCGCTGAGGCGG 4231
 Qy 432 GGAGGTGCTTCCGCGACGTGAACCCGAAAGACCGGCGGACTTTCATCGTGGCTGCTTCAC 491
 Db 4230 CCAGTTGCGGGCGGCGTGCATCCCGCAGCAGCTGCGCGTTCCTGTTGCGCGCTTCTT 4171
 Qy 492 CGGGTCCAGGCGGTCTCCGGGTACCTCCAGCGCGGAGGACCTGCGGCGACCGGATCTC 551
 Db 4170 CGGCATGACAGCAGCTCTCCGACAATCTGCACAGCGAGCGGACATCATGAGCGGTGGCA 4111
 Qy 552 GGTGATGTGGAACACGCTGCT 572
 Db 4110 GGAGCTGCGGGAGCTGATGTT 4090

RESULT 10
 US-10-156-761-3690
 ; Sequence 3690, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 3690
 ; LENGTH: 861
 ; TYPE: DNA
 ; ORGANISM: Streptomyces avermitilis
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(861)
 US-10-156-761-3690
 Query Match 11.2%; Score 78.6; DB 15; Length 861;
 Best Local Similarity 58.2%; Pred. No. 6.3e-11;
 Matches 138; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
 28 GTGCGACAGGAACGGGCGCTCGCGAGGCGGCGATCGTGGCGGCGAGCCCGCTCGGTTC 87
 13 GTGAGAGAGAGAGATCCGCGGCGGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCTC 72

QY 88 TTCAGCAGTACGGGTTCAGAGCGCCGACACAGTGGCAGAGATCTCTCGCGGGGCTCGGTTC 147
Db 73 TTCGCCCGCAGCGGTACGAAAGTGGTCTCGCTTCGACGATCAGCCGACGAGCGGGGTC 132
QY 148 ACCAAGGGCGGATGATCTTCCATCTTCGCTTCCAAAGGAGAGCTGCGCGCGCGGTGCTG 207
Db 133 AGCAACGGTGGGTGACATCTTCCATTTCCCGAGCAAGCGCCCTTGGCGGAGGCGGTGCGG 192
QY 208 GCGCAGCAGACCTTGACAGTGGCGGTGCGCGAATCCGGTCCAAAGCGGAGGAATG 264
Db 193 GTGGCGCGCGCAGCGATTTCGGGGCGGATCATCGCGCGGAGGCGGTGCGCGGCTCAG 249

RESULT 11

US-09-825-288A-1
; Sequence 1, Application US/09825288A
; Publication No. US20020192822A1
; GENERAL INFORMATION:
; APPLICANT: LEOPARDI, ROSARIO
; APPLICANT: ROIZMAN, BERNARD
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS
; FILE REFERENCE: ARCD:317USC1
; CURRENT APPLICATION NUMBER: US/09/825,288A
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 09/259,821
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 08/690,473
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4257
; TYPE: DNA
; ORGANISM: HERPES VIRUS, TYPE 1
US-09-825-288A-1

Query Match 11.2%; Score 78.2; DB 9; Length 4257;
Best Local Similarity 45.0%; Pred. No. 7.2e-11;
Matches 293; Conservative 0; Mismatches 358; Indels 0; Gaps 0;

QY 43 GCGGTCCGACGCGCGCAGCGGATCGTGGGGCAGCGCGCTCGGTCTTCGACGAGTACGGG 102
Db 1867 GCGGTCCGCGCGCGCGCAGCGGCGGAGCGCGGTGCGCGCGGTACGGCGCGCGGG 1926
QY 103 TTGAGGCGCCACATGAGCAGATCTCTTCGCGGGCTTCGATCCAAAGGCGCGATG 162
Db 1927 ATCTTCGCGCGCTGGGGCGGTGTCTCGCGCGCGCGCTCCCGCGGGGCGACGAC 1986
QY 163 TACTTCCACTTCGCTTCCAAAGGAGAGTGGCGCGCGGTGCTGGCGGAGACCTG 222
Db 1987 CCGGACGCGCGCGCGCACGCGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2046
QY 223 CAGTGGCGGTGCGGAAATCCGCTCCAGGGCGAGGACTGGTAGACTTCACTATGCTG 282
Db 2047 CGGTGGCGGTGAGTGCCTGCGCGCTGCGCGGGATCTTGAGGGGTGCGCGGAGGCG 2106
QY 283 GTGCGCCACGGCATGTGACGATCCGATCTTCGCGGGCGGGACACGGGTTCGACATG 342
Db 2107 TTGACGCGGACCTGGCGGCGGTCCCGGGCTGGCGGGCGCGCGCGCGCGCGCGCG 2166
QY 343 CAGGGGGCGGTGACATCTTCGACGCGCAACCGCTTCGGCGAGTGGGGGAGCATCTGCG 402
Db 2167 CGGCGGAGGAGACCGCGGGCGCGCGCTTCGCGCGCGCGCGCGCGCGCGCGCGCG 2226
QY 403 CAGCTCTGCGGAGGACAGAGAAACGGGGGAGGTGCTTCGCGACGTGAACCGCAAAAG 462
Db 2227 CTGCGCGGTGCGGTGCGCGAGTGGCGGTCTGTCGCGGACGCGGTGCTGCTGCGCTG 2286
QY 463 ACCGCGACTTCACTGTCGCTGCTTTCACCGGGCTCCAGCGGTCTCCCGGGTCACTTC 522
Db 2287 CGGCGGAGACCTGGCGGTGGCGCGCGGACGAGGCGCGGTGCGCGCGCGCGCGCGGT 2346
QY 523 GACCGCGAGACCTCGGCGCACCGGATCTCGGTGATGTGNAACCAACGCTGTGCGCGCATC 582

Db 2347 AGCTTGGTCCCGGGGCGCTTGGGCCCTCGCGGACCCGCGCTTCGCGAGCTCC 2406
QY 583 GTGCCCGGCTTCATGCTGACCTGGAATCGAAACCGGCGAGAGCGGATCGGAAAGGTGCG 642
Db 2407 GCGCGCGCGCGCGCGCGGACCTGCTGTTTGAACAACAGAGCTTGGCGCCCTGCTGGCG 2466
QY 643 GCGCGGCGGAGGCGCGCGGAGGCTGCGAGGCTTCCGAGGCGCGCTTCGAGAC 693
Db 2467 GCGCGCGCGAGCGCACCGGACCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCG 2517

RESULT 12

US-10-156-761-2886
; Sequence 2886, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 2886
; LENGTH: 18438
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(18438)
US-10-156-761-2886

Query Match 10.0%; Score 69.6; DB 15; Length 18438;
Best Local Similarity 47.3%; Pred. No. 1.1e-08;
Matches 281; Conservative 0; Mismatches 304; Indels 9; Gaps 2;

QY 99 CGGTTTCGAGGCGCGCACAGTGGCAGAGATCTCTCGCGGGCTCGGTTCACCAAGGCGCG 158
Db 1467 CCGGCTCGCGAGCGGACACCGCGCGGACCTCTGTCGCGGACCGGTCTCCCGCGCGCG 1526
QY 159 GATGTAATTCCTTCGCTTCCAAAGAGAGTGGCGCGCGCGGTGCTGGCGCGAGCAGAC 218
Db 1527 GCTCGGTTGCTTTCACCGGCGAGGACCGACCGCGCGCGGTGTCGCGAGCTGTA 1586
QY 219 CTGACAGTGGCGGTGCGGATCCGGCTTCCAAAGGCGAGGACTGGTAGACTTCCACT 278
Db 1587 CCGAGCTCGCGGAGTACG---CCGCCCGCTTGGACGAGGTGTCGCGAGCTTGACCC 1643
QY 279 GCTGTTGCGCCCGGATCTGCACGATCCGATCTTCGCGCGGCGGACCGGGCTCGCACT 338
Db 1644 GTGGTCCGACGCGCGCTGCGGACGTGTGTTTCGCGGCGGAGGTTCGACGAGGCGCG 1703
QY 339 GGACCAAGGGGCGGTGAGTCTTCGACCGCAACCGCTTCGCGAGTGGGGCGACATCTG 398
Db 1704 GCTCTGGACCGGTACCGAGTTTCAACCGCGCGCTTTCGCGGTGCGAGGTGCGCTCTT 1763
QY 399 CGGCCAGCTCTTGGGGGAGGACAGGAAAGGGGGAGGTGCTTCCGACGAGTGAACCGGAA 458
Db 1764 CCGGTACGCGGAGCACTGCGGCGCTGACACCGGGTTCCTGCTCGGTCACTCGGTGCTGA 1823
QY 459 AAGACCGCGGACTTTCATCTGTCGCTGCTTCAACCGGGCTCCAGGCGGTCTCCCGGGTCA 518
Db 1824 ACTGGGCGCGCGACGCTGCGCGAGTCTCTCTCTGGGGAGCGCTGACGCGTCTGCTGCG 1883

RESULT 15

US-10-282-122A-25841
; Sequence 25841, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25841
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Mycobacterium avium
US-10-282-122A-25841

Query Match 9.8%; Score 68.2; DB 17; Length 1224;
Best Local Similarity 46.4%; Pred. No. 3.1e-08;
Matches 306; Conservative 0; Mismatches 338; Indels 15; Gaps 2;
QY 36 GGNACGGGCGTCGAGCGGCGGAGCGGATCGTGGCGGAGCGCGCTCGGTCTTCACGA 95
DB 522 GCAGCGGGTGTCTCCCGGCGGCGGATTCGCGGCTGCGCGGCGGCGGAGTGTGCGGT 581
QY 96 GTACGGGTTCGAGCGGCGGCGGAGATCTCTCTCGCGGCGCTCGGTACCAAGG 155
DB 582 CCGGTGCGGAGGCGCATCACCTCGTGACCGCGGCGGCGGCGGCGGCGGCGGTG 641
QY 156 CGGATGTATTCTCACTTCGCTTCCAAAGGAAGCTGGCGCGGCGGTGTGTCGCGGAGCA 215
DB 642 GCACCTCGGATTCACGACGAGCGGCGGTTGATGTTGTTTCGCGGCGTTCGAGGCGC 701
QY 216 GACCTGACAGTGGCGGTGCGGAATCGGCTCCAAAGGCGGAGAACTGTAGACTCTAC 275
DB 702 GGCTTCGCTGAACCGGCGGCGGTGTCCGCGCGCGCGGCGGCGGCGGCGGCGGTGTC 761
QY 276 CATGCTGGTCGCCCGCGGATGTCGACGATCCGATCCGCGGCGGCGGCGGCGGCGTGC 335
DB 762 CGTGTCTGCACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 821

QY 336 ACTGGACACGAGGCGGCTGGACTTCTCCGAC-----GCCAACCCGTTCCGCGA 383
DB 822 CCGCGCCTACGTGGCGGTGCCCTACCTGGACCGAATGGACCTGACGCGCGCGCGA 881
QY 384 GTGGGCGGACATCTGGCGCCAGCTCCTGGCGGAGGACAGGAACGGGGGAGTGTCTCC 443
DB 882 CTTGGTGATCTGCCGCTCCGGGCGGATGACGGTTCGCGGAGGTGTCCGCGGTGGCC 941
QY 444 GCACGTGAACCCGMAAAGACCGGCGACTTCTCGTGGCTGCTTCAACGGGCTCCAGGC 503
DB 942 GGCCATCTAGTGGCGCTGCCGATCGGCAACGGCGAAGCGGCTCAACGGCTGCGGT 1001
QY 504 GGTCTCCGGGTCACTCTCCGACCGCGGAGACTTCGGCCACCGGATCTCGGTGATGTGAA 563
DB 1002 GGTCAACGCGGCGGCGGCGCATGGTGGTTCGCGGACGCGGCTTACGCGCGAGTTGGTCG 1061
QY 564 CCACGTGCTGCCGACGATCGTGCCG---CGTCCATGCTGACCTGGATCGAAGCGCGA 620
DB 1062 CCGGAGGTGGCGGCGCTGGTGGCGACCGCGCGGCTGGCGGCGATGACGAGCGCGC 1121
QY 621 GGAGCGGATCGGGAAGGTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 679
DB 1122 CGCGCGGTTCGGGCGATCCCGACGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCG 1180

Search completed: March 6, 2005, 22:27:00
Job time : 657 secs

This Page Blank (uspto)

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	69.2	9.9	732	8	BZ564850		BZ564850 pacs2-164
C	2	68.4	9.8	925	9	CNS0091P	AL053013 Drosophil
	3	67.2	9.6	967	8	BZ567730	BZ567730 pacs2-164
	4	66	9.4	1418	3	AK019388	AK019388 Mus muscu
C	5	66	9.3	783	9	CG450596	CG450596 OGVGU52TV
	6	65	9.3	822	9	CG444666	CG444666 OGVGE35TH
	7	64.6	9.2	925	9	CNS0091P	AL053013 Drosophil
8	63.8	9.1	1016	7	CF583098		CF583098 AGENCOURT
9	63.6	9.1	846	7	CNI46323		CNI46323 WOUND1_39
10	60.8	8.7	584	7	CK626160		CK626160 mjl18902.Y
11	60	8.6	635	7	CNS011285		CNS011285 WHE3882.C
C	12	60	8.6	935	9	CNS0065XK	AL066051 Drosophil
C	13	59.8	8.6	1084	7	CK213336	CK213336 FGAS02524
C	14	59.2	8.5	483	8	AZ212924	AZ212924 Sheared.D
C	15	59	8.4	640	2	BE363278	BE363278 WS1_61.D0
C	16	58.6	8.4	2332	9	AG363333	AG363333 Mus muscu
C	17	58.4	8.4	876	9	CG43671	CG43671 CGNAA29TH
C	18	58.2	8.3	524	9	CG3256542	CG3256542 QG1CJ26TH
C	19	58.2	8.3	744	9	CG322461	CG322461 QG2AZ31TH
C	20	58.2	8.3	819	8	BZ649366	BZ649366 OGCBY29PC
C	21	58.2	8.3	900	9	CC682633	CC682633 CGOAK35TV
C	22	58	8.3	899	9	CC620156	CC620156 QGOUBL90TV
C	23	58	8.3	932	9	CNS0072Q	AL066742 Drosophil
C	24	57.6	8.2	752	7	CK432635	UI-D-GC1-CK432635

QY

[illegible]

Db 201 CCGGCGGCGAGGTACCGTCCGGGGTCTCGACGACAGCGGTGCGGTGGCGAGATTC 260
QY 332 TCGCACTGGACAGGCGGCGGTGGACTTCTTCGACGCCAACCCGTTTCGGCGAGTGGGCG 391
Db 261 TCTCGGCGGCGGAGGAGTCTTCGACTACGAGCAAGTACACGCGCGCGGTGGCG 320
QY 392 ACATCTGCCGCCAGCTCTCGGCGGAGGACAGGAGCGGGGAGGTCTTCGACAGTGA 451
Db 321 AGGTGTTCCTCCCGACCTGACGCGGCGGCGATCGCCCGGAGGCCAGCGTCTGGCGCTGA 380
QY 452 ACCCGAAAAGACCGGCGACTTCATCTCGGTGCTTTCACCGGGTCTCAGGCGGTCTCCC 511
Db 381 AGGTACACCGGCGGTGAGCTGAGCGGTACAGCGCACCGACTTCGACTCGACGAAC 440
QY 512 GGGTCACTCCGACGCCAGGACTCTCGGCGACCGGATCTCGGTGATGTGGAACACGTCGC 571
Db 441 AGGCGCGGTCTGTCTGCTGAGGTCAATACCTCTCGGCGCATGCCCGCACCGCCTGC 500
QY 572 TGCCAGCATCGTCCGCGGCTCATGCTGATCGATCGATCGAAACCGCGCGAG 621
Db 501 TGCCGACCGCGGCGGCGCGGCGATCGAGTGTGCGGAATCTTCGCGAG 550

RESULT 2
CNS0091P/c 925 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL053013.1 GI:4934461
VERSION
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 925)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN
Query Match 9.8%; Score 68.4; DB 9; Length 925;
Best Local Similarity 14.1%; Pred. No. 0.0005;
Matches 55; Conservative 183; Mismatches 152; Indels 0; Gaps 0;

QY 168 CCATTTGCTTCCAAAGAGAGTGGCCCGCGGTGCTGGCGAGACGACCTGACGCT 227
Db 919 CSCSSBSBSCSSSMSTSSNSBSCSSBSSSTSSMSSBSSBSSSSSSSSSSSSSS 860
QY 228 GGCGGTCCCGGATCCGGTCCCAAGCGCAGGAACTGGTAGACTCACCATGCTGTGTCGC 287
Db 859 ACVKCNASSCGCGCGCMABCCMSCSSSSSCGASARGVKVRASGGAGKRGSGGASAS 800
QY 288 CCACGGCATGTCACCATCCGATCCTGCGGGGCGGACGCGGTCTCCATCTGACACGAGG 347
Db 799 HSSSSACBSSSSSSCSASCSWASSSSSSSSRSSRSGGAGGSSSSSSSSSSASAGSV 740
QY 348 GGCGGTGGAATTTCCGACGCCAACCCGTTTCGGCGAGTGGGCGACATCTCGGCCAGCT 407
Db 739 VSSASSSSSSSSSSSVSSVSSMSSMSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 680
QY 408 CTTGCGGAGGACAGGAAACGGGGGAGGTCTTCCGACGCTGAACCCGAAAGAACCGG 467
Db 679 CSTSASMSAARSSSGG 620
QY 468 CGACTTCATGTCGGTCTTACCGGGTCCAGGCGGTCTCCGGGTACCTCCGACCG 527
Db 619 GSGSVSASSGWSVSSVSSGGRSSGSGGGVGGSSSSSSSSSSSSSSSSSSSSGVCSCSSGCMCR 560
QY 528 CCAGGACCTCGGCCACCGGATCTCGGTGAT 557
Db 559 CSCSSAAAAASCVAASCMMCGKSKSGCT 530

RESULT 3
BZ567330 967 bp DNA linear GSS 17-DEC-2002
LOCUS pacs2-164_6885.Y2 pacs2-164 Pseudomonas aeruginosa genomic clone
DEFINITION pacs2-164_6885, genomic survey sequence.
ACCESSION BZ567330.1 GI:27198447
VERSION
KEYWORDS Pseudomonas aeruginosa
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.

REFERENCE 1 (bases 1 to 967)
AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V.
TITLE Whole-Genome-Sequence Variation among Multiple Isolates of Pseudomonas aeruginosa
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES
source
1..967
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pacs2-164_6885"
/clone_lib="pacs2-164"
/note="clinical isolate 2-164 Whole genomic shotgun library."

ORIGIN
Query Match 9.6%; Score 67.2; DB 8; Length 967;
Best Local Similarity 46.9%; Pred. No. 0.00088;
Matches 210; Conservative 0; Mismatches 238; Indels 0; Gaps 0;
QY 212 AGCAGACCTTGCACGTGGCGGTGCGCGATCCGCTCCAGGCGGAGGAACTGGTAGACC 271


```
Query Match 9.4%; Score 66; DB 3; Length 1418;
Best Local Similarity 47.0%; Pred. No. 0.0015;
Matches 259; Conservative 0; Mismatches 275; Indels 8; Gaps 2;

QY 143 CGGTACCAAGGCGGCGATGTTTCACATTCCTTCCAAAGGAAGAGTGGCCCGCGCG 202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 CGGTACCGGACTCGTGGTGGTGAACCCCGAGACGAAGCGCTCCCGCTGCTTCGGCT 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 203 TGTGCGCGAGACCCCTGCAGTGGCGGTGCGGAATCCGGCTCAAGGCGCAGAAC 262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 TGTGACCTACTCGAACGTTGAGAGCGGATGCCGCCATGGCCGCTGCGCCGCGCG 482

QY 263 TGTAGACCTCCACATGCTGTCGCCACAGCATGCTGCACATCCGATCCCTCGCGCGCG 322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 TGAACGGCAACACGCTGGAGCTGAAGGCGCGCTGTCGCGGAGGATTCGGCGCGCGCG 542

QY 323 GCACCGGCTCGACCTGGACACAGGGCGGCTGACATTCCTCCGACGCCAACCCGCTTCGGCG 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 543 GGGCGCACGCAAGGTGAAGAAGCTGTCGTGGCGCGCTCAAGGGCGAC---GTGGCGG 599

QY 383 AGTGGGCGGACATCTGGCCCGACCTCTTCCGCGAGGACAGGAAAGGGGAGGTGCTTC 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 600 AGGGACACCTGATTCGAGACTTCTCGCAGTTCGGCGCGGTGGAGAAAGGCGGAGATCAT 659

QY 443 CGCACGTGAACCCGAAAAAGACCGGCGACTTTCATCGTGGCTGCTCTTCCCGGCTCCAGG 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 660 CCGACAAGCAGTGGGCAAGAGCGCGCTTCGGCTTCGCTTCTTCCAGAGCCAGACG 719

QY 503 CGGTCTCCCGGTGTCACCTCGACCGCGCAGGACCTCGGCCACCGGATTCGGTGAATGGA 562
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 720 CGGCCGACAAAGGCGCGGTGGTCAAGTTCCACCGGATCCAGGGCCACCGCGTGGAGTGA 779

QY 563 ACCACGTGTCGCCACGATGTCGCCGCGTCCATGCTGACCTGATCGAAGACCGGCGAGG 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 780 AGNAGCGGTGCCAAG-----GAGATATCCACGCGGCGGCGGGGTGCGCGGCGCG 834

QY 623 AGCGGATCGGGAAGTTCGCGCGCGCGCGCGAGCGCCCGAGGCTGCGGAGGCTTCGAGG 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 835 CCGGGCGGCGCGCGGCGGAGCGCGGCGCGCGCGCGGCGGCGCGCGCGCGCGCGAGA 894

QY 683 CC 684
    |||
Db 895 CC 896

RESULT 5
CG450596
LOCUS
DEFINITION
  OGVG52TV ZM 0.7_1.5 KB Zea mays genomic clone ZMMBma0479107,
  genomic survey sequence.
ACCESSION
  CG450596
VERSION
  CG450596.1 GI:34835596
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 783)
  Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
  Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
  Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
  Consortium for Maize Genomics
  Unpublished (2002)
  Other GSs: OGVG52TH
  Contact: Cathy Whitelaw
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TP
  Class: sheared ends.

CG450596 783 bp DNA linear GSS 17-SEP-2003
OGVG52TV ZM 0.7_1.5 KB Zea mays genomic clone ZMMBma0479107,
genomic survey sequence.
ACCESSION
  CG450596
VERSION
  CG450596.1 GI:34835596
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 783)
  Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
  Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
  Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
  Consortium for Maize Genomics
  Unpublished (2002)
  Other GSs: OGVG52TH
  Contact: Cathy Whitelaw
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TP
  Class: sheared ends.
```

FEATURES
Source

```
Location/Qualifiers
1..783
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0479107"
/clone_lib="ZM 0.7_1.5 KB"
/note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
```

ORIGIN

```
Query Match 9.3%; Score 65; DB 9; Length 783;
Best Local Similarity 47.0%; Pred. No. 0.0025;
Matches 240; Conservative 0; Mismatches 265; Indels 6; Gaps 1;

QY 54 GCGGAGGCGATGTCGCGGCGAGCCGCTTCGACAGTAGTACGGGTTTCAGGCCCG 113
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 208 GCAGCAGCGCGCGTGGGTGCTCGCTGCTGCCCTAGCTTCGATCGCGCCCTACGC 267

QY 114 GACAGTGGCAGATCTCTTCGCGGCGCTCGGTACCAAGGCGCGATGTACTTCACATT 173
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 268 CACCATGCACGACGAGCTGTCGCGCTGCGCTCGAGAAGCGCGCATGCTCATGTCGT 327

QY 174 CGCTTCCAAAGGAAGAGCTGGCCGCGGTGCTGGCCGAGCAGACCTGCACGTGGCGGT 233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 328 GCCTTCCACACGCGCTCGCATCGACGGCTTCGTCGAGAACAACGCGCAACGCCGG 387

QY 234 GCGGAAATCCGGCTCAAGGCGCGAGAACTGCTAGACCTTCACCATGCTGGTGGTCCGCG 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 388 CTCCTCCAGCGCGCAACGTCACGCTCCACTACTCCCTG-----CTCGCTGCG 441

QY 294 CATGCTGCACCATCCGATCCTCGCGGCGGCGACGCGCTCGACCTGGACCGAGGCGCGT 353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 CATCTCTGTCGACCGGGGAGCCTCTCTGTCGATGTCGCGCGGTGGCGCGGCGCGCGCG 501

QY 354 GCACCTTCTCCACGCGCAACCGCTTCGCGGAGTGGGCGACATCTCGCGCCAGCTCCTCGCG 413
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 502 GSCCAGACGCGCGCGAGCGCTTCCTCACCAGCGTTCGCTCTACTTCTTCGCGCGGCG 561

QY 414 GAGGCGCAGGAAACGCGGCGAGGTCTTCGCGACGTGAACCCGAAAAAGACCGCGGACATT 473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 562 GGACGACCGGAGCGCTCGCGCTCGCCACGCACATGGCCGAGGACGCGCGCCATCGGCT 621

QY 474 CATCTGCGGTGCTTTCACCGGCTCCAGCGGTCTCCCGGTCACCTCCGACCGCAGGA 533
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 622 CACCTCTTCGCTTCCTGCGCGCGCGAGTGGCGCAAGGCGCGGCGCGCGCGAGGAGA 581

QY 534 CTTCCGCGCCACCGGATCTCGGTGATGTGAAC 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 682 CCGGCTCGACGAGGAGCGCTGCAGAGGTAC 712

RESULT 6
CG444666/c
LOCUS
DEFINITION
  OGVGE35TH ZM 0.7_1.5 KB Zea mays genomic clone ZMMBma0475E21,
  genomic survey sequence.
ACCESSION
  CG444666
VERSION
  CG444666.1 GI:34826493
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 822)
  Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
  Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
  Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
  Consortium for Maize Genomics
  Unpublished (2002)
  Other GSs: OGVGE35TV
```


Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES

Location/Qualifiers
1..822
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0475E21"
/clone_lib="ZM 0.7 1.5 kb"
/note="Vector: pBCKS-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 9.3%; Score 65; DB 9; Length 822;
Best Local Similarity 47.0%; Pred. No. 0.0025;
Matches 240; Conservative 0; Mismatches 265; Indels 6; Gaps 1;
QY 54 GCGGAGGCGATCGTGGGGAGCGCCCTCGCTTCGACGAGTACGGGTTGAGCGCG 113
DB 525 GCGAGCGCGCGCGTGGGGTGGCGGTGCTGCTCCCTACGTCATCGCGCCCTACGC 466
QY 114 CACAGTGGCAGAGATCTCTCGGGGCGCTCGGTCAACAGGGCGCATGACTTCCACTT 173
DB 465 CACCATGACGACGAGCTCTGGCGCTGGCGCTCGAGAGCGGCATGCTCATGCTGT 406
QY 174 CGCTTCCAGAGAGAGCTGGCGCGCGCTGTGCGCGAGCAGACCTCGACGTCGCGGT 233
DB 405 GCGCTTCCACGCGCGCTCGCATCGACGCTCGGTGCGAGAACACGACGCGCAGCGCG 346
QY 234 GCGGNAATCGGCTCAAGGCGCAGAACTGGTAGACCTCACCATGCTGGTGGCCACGG 293
DB 345 CTCGCTCCAGGCGCGCAACGTCAACGTGTCCACTACTCCCGCTG-----CTCGTCCG 292
QY 294 CATGCTGACGATCCGATCTCGGGCGGCGACGCGGCTCGCACTCGACGAGGGGCGGT 353
DB 291 CATCTCTGTCGACCGGGGAGCTCTCGTCGATGCTCGCGGTGGCGGGGCGCCGC 232
QY 354 GGAATTTCTCGAGCGCAACCCGTTGCGCGAGTGGGCGCAGATCTGCGCCAGCTCTGCG 413
DB 231 GGCACAGCGCGCGCGAGCTTCCCTCACGCGTGGCGCTACTTCTCGCGGGCGC 172
QY 414 GGAGGACAGGAACGGGGGAGGTGCTTCGCGACGTGAACCCGAAAAAGACCGGCGACTT 473
DB 171 GGACGACCGGGAGCGCTCGCGCTCGCCACGACATGGCGGAGCGCCCATCGGCT 112
QY 474 CATGTCGCTGTTCTCACCGGGTCCAGCGGGTCTCCCGGTCCACCTCGACCGCCAGGA 533
DB 111 CACCGTCTTCGGTTCTCGCGCGCGGAGTGGCGCAAGGGCGGCGACCCGAGAGGA 52
QY 534 CCTCGGCCACCGGATCTCGGTGATGTGAAC 564
DB 51 CCGGCTCGAGGAGGCGCTGCGAGGAGTAC 21

RESULT 7

CNS0091P 925 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR19D16 of RPI-98 library from Drosophila melanogaster. (fruit fly), genomic survey sequence.

ACCESSION

AL053013 GI:4934461

VERSION

AL053013.1 (fruit fly)

KEYWORDS

Drosophila melanogaster

SOURCE

Drosophila melanogaster

ORGANISM

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

REFERENCE

AUTHORS
TITLE
JOURNAL

COMMENT

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 925)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Oosawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw ap, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Location/Qualifiers
1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPI-98"
/note="end : TET3"

ORIGIN

Query Match 9.2%; Score 64.6; DB 9; Length 925;
Best Local Similarity 14.3%; Pred. No. 0.003;
Matches 53; Conservative 183; Mismatches 132; Indels 3; Gaps 1;
QY 332 TCGCTGACGACGAGGGCGGTGGACTTCTCGACGCCAACCGTTGCGGAGTGGGCG 391
DB 554 TSSGGYGGKSGSGBSGCCSCSCSSSCSCBCCCCSCSYSCSSSSSSKCSST 613
QY 392 ACATCTGCCCGCAGCTCTCGCGGAGGACAGACGCGGAGGTCTTCCGACGCTGA 451
DB 614 SBCSCCCKSVKVCGTSCSSSSSSSSSTSSSTSSKSSSSSSSSSSSYTTSK 673
QY 452 ACCCGAAAAGACCGCGACTTTCATGTCGCTGTTCACCGGCTCCAG---GCGGTCT 508
DB 674 STSAGSGSWAGGGGSGTGSTSSSSSSSTSTSSSVSSGSKSSTBSGSSSSGSSS 733
QY 509 CCGGGTACCTCCGACCGCAGGACCTCGGCACCGGATCTCGGTGATGTGAACACG 568
DB 734 SSTSSBBSCTSTSSSSSSSYSTSCCTCCCSYSSSTSSSTSSSTSSSTSSSVG 793
QY 569 TCTGCTCCAGCATGTCGCGCGCTCCATGCTGACCTGGATCGAAACCGCGAGAGCGGA 628
DB 794 TSSSDSTSTCCSCCYMTCTCCYBMBCTTSTCGSSSSSSGKGGVTTCGCGGCGSSST 853
QY 629 TCGGAAGGTTCGCGCGCGCGGAGCGCCGAGGCTCGGAGGCTCGAGGCGCGCT 688
DB 854 NGBMGTSSACSSSSSSSVSSSKSSASSSSVSSGSSGSSVSSSSASKSSSSGVS 913
QY 689 CCGACGAGTAG 699
DB 914 SGSGSGSGSVS 924

RESULT 8

CF583098

LOCUS

DEFINITION

ACCESSION

VERSION

CF583098 1016 bp mRNA linear EST 24-SEP-2003
AGENCY: 11361399 updated NIH_MGC_137 Mus musculus cDNA clone
IMAGE: 6432215 5', mRNA sequence.
CF583098
CF583098.1 GI:35196360

distribution through Research Genetics, Alabama, USA. Sheared DNA end sequences search page: <http://www.tigr.org/tdb/mdb/tbdb/>.
Seq primer: M13-Reverse
Class: shotgun.

FEATURES
source Location/Qualifiers
1..483
/organism="Trypanosoma brucei"
/mol_type="genomic DNA"
/strain="TREU927/4 Gutat 10.1"
/db_xref="taxon:5691"
/clone="Sheared DNA-112H10"
/clone_lib="Sheared DNA"
/note="Vector: pUC18; Site 1: SmaI; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 Gutat 10.1) was mechanically sheared to give a tight size distribution (approx 2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Borell, Oxford University Press, 1999)."

ORIGIN

Query Match 8.5%; Score 59.2; DB 8; Length 483;
Best Local Similarity 49.5%; Pred. No. 0.04;
Matches 213; Conservative 0; Mismatches 208; Indels 9; Gaps 2;
QY 265 GTAGACCTCACCATGTCGTCGCCACGCGATGTCGACATCCGATCTCTGGGGGGGC 324
Db |||||
QY 454 GTAGGAGCAACCTGTCGTCGCCGACGCGCGGTGATCATCCGCTGGGC 395
QY 325 ACGCGGCTCGCACTGACACGAGGGGGGTGGACTTCTCCGACGCCAACCCGTTCCGCGAG 384
Db |||||
QY 394 AAGGGCTTCAACACGTCGAGCCCTGGCGACACCGTATCAAGCCGGCGAGCCGTT 335
QY 385 TGGGGCGCATCTGCGCCCGAGTCTCTGGCGAGGACACAGGAACGGGGGAGTGTTCG 444
Db |||||
QY 334 CCGGACGCCATCTGCGCCGCAAGCGCGCGAGGGGGGTGCGGGGTGGAATCTAT 275
QY 445 CACGTGAACCCGAAAGACCGCGACT-----TCATCGTGGTGTTCACCGGGCTC 498
Db |||||
QY 274 GCGCGCATCCCGGACGATCGCGCGGGGTGATGAACCGCGGTGTACGGCTCG 215
QY 499 CAGGCGGTCTCCGGGTCACTCCGACCCGCA---GGACCTCGGCGCACCGGATCTCGGTG 555
Db |||||
QY 214 GAGACGTCATGTGTCAAGTCGGTGGGTGATGGATCGCGCGGTGTGCGGAG 155
QY 556 ATGTGGAACCGATGTCGCCAGCATCTGTCCGGGTCCATGTCGATCGGATCGAAACC 615
Db |||||
QY 154 CTGTCGGTGGAGAGTGCATACACCTACCGTCAAGCGGCTGCGAGGACGCGGAGACC 95
QY 616 GCGGAGGACGATCGGGAAGTCTCGCGGGCGCGGAGCGCGGCTCGGAGGCC 675
Db |||||
QY 94 GTATCGTACTGACCGCATCTTCGAGGGGGCGCGGACGAGCCCGAGGCGATCAAGGCC 35
QY 676 TCCGAGGCCG 685
Db |||||
QY 34 CGCATGGCCG 25

RESULT 15
BE363278/c
LOCUS
DEFINITION WS1_61_D03_g1_A002 Water-stressed 1 (WS1) Sorghum bicolor cDNA,
mRNA sequence.
ACCESSION BE363278
VERSION BE363278.1 GI:9304835
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.

1 (bases 1 to 640)

Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt,L.H.

An EST database from Sorghum: water-stressed plants

Unpublished (2000)

Contact: Cordonnier-Pratt MM

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.

Seq primer: PolyTMix

High quality sequence start: 50

High quality sequence stop: 574

POLYA=No.

FEATURES
source Location/Qualifiers

1..640

/organism="Sorghum bicolor"

/mol_type="mRNA"

/db_xref="taxon:4558"

/clone_lib="Water-stressed 1 (WS1)"

/note="Organ: Mix of 5-week old plants on days 7 & 8 after

water was withheld; Vector: Lambda Zap; Site 1: XhoI;

Site 2: EcoRI; The library was made from poly-A RNA in the

cloning vector lambda ZAP II. Clones to be sequenced were

prepared by mass excision."

ORIGIN

Query Match 8.4%; Score 59; DB 2; Length 640;
Best Local Similarity 49.3%; Pred. No. 0.044;
Matches 181; Conservative 0; Mismatches 185; Indels 1; Gaps 1;
QY 316 CGGCGGGCACGCGGCTCGCACTGGACACGAGGGGGCGGTGGACTTCTCCGACGCCAACCG 375
Db |||||
QY 368 CGCCGCTCTGCGCGCGCGCAGGTACACGTGCGGCGGCACCCACCGCGCGCGCGCG 309
QY 376 TTGCGCGAGTGGGGCGACATCTGCGCCCGAGTCTTGGCGGAGGACAGAACGGGGGGAG 435
Db |||||
QY 308 CTTGCGGTGCGCTCGTGTCTGCTCCCTCATCTCGAACAGCGCGGCTC 249
QY 436 GTGCTTCCGCACTGAACCCGAAAGACCGCGGACTTTCATCGTGGCTGTTCACCGGG 495
Db |||||
QY 248 CGGTCCGGGACGGGAAGCGGGCGGCCACGCGGGAACCTCGACGGGCGCGACCGCGC 189
QY 496 CTCAGGCGGTCTCCCGGGTCACTTCGACCGCCAGGACCTCGGCGACCGGATCTCGGTG 555
Db |||||
QY 188 CACTTGGCGCGCGCGCGCGCGTCCCGAACCGCTGGCTCAGCCCTCCCTGCACCGCGC 129
QY 556 ATGTGGAACACAGTGTCTCCCGAGCATCG-TGCGGCGTGCATGTGACCTGTGATCGAAC 614
Db |||||
QY 128 CCGCGGGCGCGCGCTGCGCGCGCGCTCTGCGCTCCACTCCACTGCGCGGAGCGC 69
QY 615 CGGCGAGGAGCGGATCGGGAAGTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 674
Db |||||
QY 68 CGCGGGGACGGGCTCCCGGGGTGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 9
QY 675 CTCGAG 681
Db |||||
QY 8 CTCGCGC 2

Search completed: March 6, 2005, 22:12:04
Job time : 3474 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2005, 20:07:07 ; Search time 149 Seconds
(without alignments)
602.204 Million cell updates/sec

Title: US-10-049-710A-2
Perfect score: 1164
Sequence: 1 MAVRHERVAVROERAVRTQ.....VAAAEAAAEAAEAAEAASDE 232

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1164	100.0	232	5	AAM48990 S virgini
2	476	40.9	215	5	Abb80940 S. coelic
3	476	40.9	215	7	Adel4791 Streptomy
4	434.5	37.3	222	8	Adh39751 Streptomy
5	383	32.9	80	8	Adh39814 BarA amin
6	338.5	29.1	224	3	Ray83794 S. fradia
7	243	20.9	214	4	Aau04043 Streptoco
8	236	20.3	228	8	Adh39748 Streptomy
9	233	20.0	80	8	Adh39815 FarA amin
10	224	19.2	80	8	Adh39813 ArpA amin
11	214	18.4	213	8	Adh39743 Streptomy
12	161.5	13.9	203	4	Aau04039 Streptoco
13	138.5	11.9	211	6	Ada34808 Acinetoba
14	138	11.9	195	2	Aaw42428 Glucuroni
15	135.5	11.6	200	6	Abu16923 Protein e
16	134	11.5	201	7	Abu02144 Pseudomon
17	132.5	11.4	295	6	Abu21858 Protein e
18	131.5	11.3	273	4	Abu20218 Protein e
19	125	10.7	111	4	Abbs3089 Escherich
20	124.5	10.7	253	8	Adi39286 S. hygrois
21	124.5	10.7	313	4	Abg30006 Novel hum
22	122	10.5	295	5	Abba48122 Listeria
23	120.5	10.4	188	4	Aag91446 C. glutami
24	120.5	10.4	188	7	Add13970 C. glutam
25	120.5	10.4	202	5	Abb49955 Listeria

26	120	10.3	205	6	ABU21324	Abu21324 Protein e
27	119	10.2	229	7	ABO63823	Abos3823 Klebsiell
28	118.5	10.2	81	6	ABU22851	Abu22851 Protein e
29	118.5	10.2	216	7	ABO71378	Abos71378 Pseudomon
30	118	10.1	280	6	ABR55197	AbR55197 Amino aci
31	117.5	10.1	231	6	ABM69166	Abm69166 Photorhab
32	115.5	9.9	189	4	ABBS2817	Abbs2817 Escherich
33	115.5	9.9	230	7	ABO67178	Abos7178 Klebsiell
34	115.5	9.9	234	7	ADF04427	Adf04427 Bacterial
35	115	9.9	182	6	ABU34193	Abu34193 Protein e
36	115	9.9	254	7	ABO69312	Abos69312 Pseudomon
37	114	9.8	194	6	ABU19797	Abu19797 Protein e
38	114	9.8	212	6	ABJ18817	Abj18817 Pseudomon
39	114	9.8	213	8	ADQ03074	Adq03074 P. aerugi
40	114	9.8	230	7	ABO73318	Abos73318 Pseudomon
41	113.5	9.8	213	7	ABO79293	Abos79293 Pseudomon
42	113	9.7	199	6	ADB12647	Adb12647 Alloiococ
43	113	9.7	214	6	ABM70661	Abm70661 Photorhab
44	113	9.7	243	7	ABO78019	Abos78019 Pseudomon
45	112.5	9.7	235	7	ABO67541	Abos67541 Klebsiell

ALIGNMENTS

RESULT 1
AAM48990
ID AAM48990 standard; protein; 232 AA.

XX AAM48990;
XX 03-MAY-2002 (first entry)
XX S virginiiae butanolide binding protein.
XX Butanolide binding protein; gene expression induction; operator;
XX transgenic plant; antibiotic production.
XX Streptomyces virginiiae.

XX WO200196581-A1.
XX 20-DEC-2001.
XX 15-JUN-2001; 2001WO-JP005096.
XX 15-JUN-2000; 2000JP-00180466.
XX (KANF) KANEKA CORP.
XX Shimmyo A, Kato K, Yamada Y, Nihira T, Shindo T;
WPI: 2002-098073/13.
XX N-PSDB; AAL44719.
XX Inducing expression of gene under regulation by operator at actinomycetes
XX self-regulator provision site, useful in producing transformant tobacco
XX for production of antibiotics e.g. virginiamycin.
XX Disclosure; Page 48-49; 57pp; Japanese.

XX The present invention relates to a method of inducing the expression of a
XX gene under the regulation of an operator at an actinomycetes self-
XX regulator provision site comprising imparting the characters of a
XX repressor and an operator constituting a gene expression-inducing system
XX with use of the actinomycetes self-regulator as inducer to a plant by
XX gene transfer thereby providing the self-regulator to the transformed
XX plant. The method is useful in producing transformant tobacco for
XX production of antibiotics e.g. virginiamycin. The present sequence is the
XX Streptomyces virginiiae butanolide binding protein which was used in the
XX exemplification of the invention
XX Sequence 232 AA;

	Query Match	100.0%;	Score 1164;	DB 5;	Length 232;
	Best Local Similarity	100.0%;	Pred. 0.34e-120;		
	Matches 232;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
1	MAVRHERVAVRGERAVRTRQAIIVRAAASVFDYGFEEATVAEITLSRASVTKGAMTFHFAAS	60			
1	MAVRHERVAVRGERAVRTRQAIIVRAAASVFDYGFEEATVAEITLSRASVTKGAMTFHFAAS	60			
61	KEELARGVLAEOQLTHVAVPESGSKAQELVDLTWLVVAHGMLHDPPIIRAGTRRLALDQGVDF	120			
61	KEELARGVLAEOQLTHVAVPESGSKAQELVDLTWLVVAHGMLHDPPIIRAGTRRLALDQGVDF	120			
121	SDANPFGEWGDDICAOQLLAEAQERGEVLPVHNPKTKGDFIVGCGFTGLQAVSRVTSORQDLG	180			
121	SDANPFGEWGDDICAOQLLAEAQERGEVLPVHNPKTKGDFIVGCGFTGLQAVSRVTSORQDLG	180			
181	HRISVMNHNVLPSIVPASMLTIETGEERIGKVAIAAAAEAAAEAAAEAAAEAAAEAAASDE	232			
181	HRISVMNHNVLPSIVPASMLTIETGEERIGKVAIAAAAEAAAEAAAEAAAEAAAEAAASDE	232			

RESULT 2
BB80940
ABB80940 standard; protein; 215 AA.
ABB80940;
21-OCT-2002 (first entry)
S. coelicolor Scbr protein.
Antibiotic; bacterium; scbA; afsA; scbR; arpA; barA; actinorhodine; Act;
undecylprodigiosin; Red.
Streptomyces coelicolor.
CA2322241-A1.
23-APR-2002.
23-OCT-2000; 2000CA-02322241.
23-OCT-2000; 2000CA-02322241.
(PLAN-) PLANT BIOSCIENCE LTD.
Takano E, Bibb M;
WPI; 2002-501089/54.
N-PSDB; ABN86379.
Modifying antibiotic-producing Streptomyces, to increase, or alter timing
of, antibiotic production, by deleting the scbA or scbR genes.
Claim 20; Fig 9; 64pp; English.

strain or heterologues to increase production of antibiotics or to alter the timing of antibiotic production. The modification is functional deletion of the *scbA* gene of *S. coelicolor*, or its homologues, but is not deletion of the *afsA* gene of *S. griseus*, or the modification is functional deletion of the *scbR* gene of *S. coelicolor*, or its homologues, but is not deletion of *arpa* of *S. griseus* nor *barA* of *S. virginiae*. The method is particularly used for production of the antibiotics actinorhodine (Act) and undecylprodigiosin (Red). The present sequence represents the *S. coelicolor* *ScbR* protein

Sequence 215 AA;	
Query Match	40.9%; Score 476; DB 5; Length 215;
Best Local Similarity	47.3%; Pred. No. 5.9e-44;
Matches 97; Conservative 36; Mismatches 72;	Indels 0; Gaps 0

QY	11	ROEAVRTRQAI	VRAAASVFDEY	FEAA	TVAEIIS	RSASVTR	KGAMYFHF	FASKEEL	ARGVLA	70	
		: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	: : : :		
DB	3	KQDRAIR	TRQITL	DAQAQV	FEKQGYQA	ATITEIL	KVAGVTK	GALYFHF	QSKKEAL	GVFD 62	
		: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	: : : :		
QY	71	EOTLHV	APESGSKA	QELAVDL	TM	LVAHG	MLDPTIL	RAGTR	LALD	OGAVDFSDANP	GEWG 130
		: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	
DB	63	AQEP	QAVPEQ	PLRLQEL	UIDM	GMFLF	CHRLR	TNNV	VARAGV	RLSMDOQA	HGLDRRGPFRRWH 122
		: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	
QY	131	DIC	QAQLAAE	AOERGE	VLPHV	NPKTG	DFIVG	CTGLQ	AVSRVTS	DRQDLG	HRISYMMNHV 190
		: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	
DB	123	ETLLK	LNQA	KENGELL	PHV	VT	TD	SADLY	VGT	FAGIQV	VSQVSDYQDLEHRYALLQKH 182
		: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	: : : :	
QY	191	LPSIV	PASMLT	WIETGE	RI	CKVAA	215				
		: : : :	: : : :	: : : :	: : : :	: : : :	: : : :				
DB	183	LP	AI	AVPSV	LAALD	SEER	GARLAA	207			
		: : : :	: : : :	: : : :	: : : :	: : : :	: : : :				
RESULT 3											
ADE14791											
ID	ADE14791 standard; protein; 215 AA.										
XX	ADE14791;										
AC											
XX											
DT	29-JAN-2004 (first entry)										
XX											
DE	Streptomyces coelicolor Scbr protein.										
XX											
KW	antibiotic-producing strain; antibiotic production; scbA gene; Scbr gene;										
KW	actinorhodin; undecylprodigiosin.										
XX											
OS	Streptomyces coelicolor.										
XX											
FN	US2003124644-A1.										

PD	03-JUL-2003.	
XX		
FF	23-OCT-2001; 2001US-00017471.	
XX		
PR	23-OCT-2000; 2000US-0242561P.	
XX		
PA	(TAKA/) TAKANO E.	
XX	(BIBB/) BIBB M J.	
XX		
PI	Takano E, Bibb MJ;	
XX		
DR	WPI; 2003-810983/76.	
DR	N-ESDE; ADE14790, ADE14794.	
XX		
PT	Modifying an antibiotic-producing strain of Streptomyces coelicolor or	
PT	Streptomyces lividans to increase or alter the timing of antibiotic	
PT	production in the strain, comprises functionally deleting in the strain	
PT	the scbA or ScbR gene.	
XX		
ES	Claim 28; Fig 9; 33pp; English.	
XX		
CC	This invention relates to the novel modification of an antibiotic-	
CC	producing strain of Streptomyces coelicolor or Streptomyces lividans to	
CC	increase or to alter the timing of antibiotic production in the strain.	
CC	The method comprises functionally deleting in the strain the scbA or ScbR	
CC	gene. The method is useful in increasing and altering the timing of	
CC	antibiotic production (especially actinorhodin and undecylprodigiosin) in	
CC	Streptomyces species, particularly Streptomyces coelicolor or	
CC	Streptomyces lividans. The present sequence is that of the Streptomyces	
CC	coelicolor ScbR protein which is used in the method of the invention.	
XX		
SQ	Sequence 215 AA;	

Query Match	40.9%	Score 476;	DB 7;	Length 215;
Best Local Similarity	47.3%;	Pred. No. 5.9e-44;		
Matches	97;	Conservative	36;	Mismatches 72; Indels 0; Gaps 0
Qv	11	QREAVRTRQAIIVRAAASVDEYGFEEATVAEILSRASVTKGAMVFHPFASKEELARGVLA	70	

QY	69	LAEOQLHVAVPESGSKAQEL	88
Db	61	LAEOQLHVAVPESGSKAQEL	80
RESULT 6			
AAU04043	AAU04043 standard; protein; 224 AA.		
XX	AAU04043		
XX	AAU04043		
XX	AAU04043		
DT	21-JUN-2000	(first entry)	
XX	S. fradiae tylosin biosynthetic pathway regulatory protein.		
DE	Gene cluster; macrolide antibiotic; tylosin; biosynthetic pathway;		
KW	recombinant; Streptomyces; hybrid antibiotic.		
XX	Streptomyces fradiae.		
OS	Streptomyces fradiae.		
XX	WO200009168-A1.		
PN	17-FEB-2000.		
XX	02-AUG-1999;	99WO-ES000248.	
PF	05-AUG-1998;	98ES-00001692.	
XX	(ANTI) ANTIBIOTICOS SAU.		
XX	Fouces Martinez R, Mellado Duran E, Diez Garcia B;		
PI	Esteban Morales M, Bernasconi E, Barredo Fuente JL;		
XX	WPI; 2000-205718/18.		
DR	N-PSDB; AA299100, AA299101, AA299102, AA299103, AA299104.		
XX	Increasing tylosin and/or hybrid antibiotic production in Streptomyces		
PT	species by introducing genes of the Streptomyces fradiae tylosin		
PT	biosynthetic cluster.		
XX	Claim 18; Page 54-55; 63pp; Spanish.		
XX	This sequence corresponds to a regulatory protein encoded by the gene		
CC	cluster encoding proteins involved in the macrolide antibiotic tylosin		
CC	biosynthetic pathway. The invention relates to methods of increasing the		
CC	level of production of tylosin and/or hybrid antibiotics in Streptomyces		
CC	spp. comprising culturing cells that have been transformed with a vector		
CC	containing a Streptomyces fradiae genomic DNA fragment. The method is		
CC	useful either for producing recombinant S. fradiae strains that		
CC	overproduce the antibiotic tylosin or for producing recombinant strains		
CC	of other Streptomyces spp. that will produce novel hybrid antibiotics		
XX	Sequence 224 AA;		
XX	Query Match	29.1%; Score 338.5; DB 3; Length 224;	
XX	Best Local Similarity	43.6%; Pred. No. 1.1e-28;	
XX	Matches	82; Conservative 32; Mismatches 69; Indels 5; Gaps 5;	
QY	11	ROEAVRTRQAIIVRAAASVDFEYGFPAATVAEILSRASVTGKAMYFHFASKELARGVLA	70
Db	3	ROEAAQTRRTIV-AAAASFDELGYEATTTAEILKRSVTKGALYFHFASKELARGVLT	61
QY	71	EOTLHV-APVPSGSKAQELVDLMLVAHGMH- DPILRAGTTRIALDQGV-DFSDAN-PF	126
Db	62	SQRAVPPVEQRLVQQIIDEITLLAQLLSKGDPLVGRSVRLTVFPGPRDGLDRAPM	121
QY	127	GEMGDI CAQLLAQAQERGEVLPHVNPCKTGDFIVGCTFTGLQAVSRVTSRQDILGHRISVM	186
Db	122	GEMGDI CAQLLAQAQERGEVLPHVNPCKTGDFIVGCTFTGLQAVSRVTSRQDILGHRISVM	181
QY	187	WNHVLPSI 194	
		: :	

Db	182	HRHLMTSV	189
RESULT 7			
AAU04043	AAU04043 standard; protein; 214 AA.		
XX	AAU04043		
XX	AAU04043		
XX	AAU04043		
DT	11-SEP-2003	(revised)	
DT	23-OCT-2001	(first entry)	
XX	Streptococcus coelicolor Mmfr protein.		
DE	SCPI; methylenomycin cluster; mmc; MmyR; MmF; MmF; MmF; MmF; MmF; MmF;		
KW	MmyO; MmyG; MmyU; Mmr; heterologous gene expression.		
XX	Streptomyces coelicolor.		
OS	Streptomyces coelicolor.		
XX	WO200148228-A1.		
PN	05-JUL-2001.		
XX	20-DEC-2000;	200WO-GB004972.	
PF	23-DEC-1999;	99GB-00030477.	
XX	(PLAN-) PLANT BIOSCIENCE LTD.		
XX	Chater KF, Bruton CJ, O'rouke SJ, Wietzorrek AW;		
PI	WPI; 2001-425675/45.		
DR	N-PSDB; AA507627, AAU04045, AAU04046.		
XX	Novel expression cassette for expressing a nucleic acid of interest,		
PT	derived from the regulatory region of methylenomycin gene cluster of SCPI		
PT	plasmid of Streptomyces coelicolor A3(2).		
XX	Claim 25; Fig 8e; 142pp; English.		
XX	The sequence represents the Mmfr protein encoded by the mmfr gene carried		
CC	on the expression cassette present on plasmid SCPI. The expression		
CC	cassette is the regulatory region of the methylenomycin cluster (mmc)		
CC	from Streptomyces coelicolor A3(2), which encodes the MmYR, MmF, MmF,		
CC	MmF, MmF, MmY, MmY, MmY, MmY, MmY, MmY, MmY, MmY, MmY, MmY, MmY,		
CC	expression cassette is useful for expressing a nucleic acid of interest,		
CC	substantially only when the host cell culture reaches high cell density		
CC	at or close to the stationary phase of host cell culture. In particular		
CC	the system is useful in regulating methylenomycin production. Reduced or		
CC	no expression of the nucleic acid of interest is observed earlier in		
CC	growth, avoiding toxic effects of some gene products on growth and the		
CC	system does not require addition of exogenous inducer. The methylenomycin		
CC	cluster naturally present on a highly transmissible plasmid permits		
CC	properly regulated expression in diverse Streptomyces host and the		
CC	expression is driven by a strong promoter, leading to high yield of the		
CC	desired end product. (Updated on 11-SEP-2003 to standardise OS field)		
XX	Sequence 214 AA;		
XX	Query Match	20.9%; Score 243; DB 4; Length 214;	
XX	Best Local Similarity	32.6%; Pred. No. 4e-18;	
XX	Matches	59; Conservative 34; Mismatches 76; Indels 12; Gaps 3;	
QY	11	ROEAVRTRQAIIVRAAASVDFEYGFPAATVAEILSRASVTGKAMYFHFASKELARGVLA	70
Db	23	QOERSIKTRQAIILEAASEIFASRGYRGSVDVSRVGMTKGVYFHFPSKESLATAYVE	82
QY	71	EOTLHVAVPESGSKAQELVDLMLVAHGMH- DPILRAGTTRIALDQGVDFSS 121	
Db	83	E---HYARWPAAMEEIRIQGFTFLETVEEMLHRAAQAQFDDPVMQAGARLQSERAFIDAE	139
QY	122	DANPFGEMGDICQAQLLAQAQERGEVLPHVNPCKTGDFIVGCTFTGLQAVSRVTSRQDILGH	181
		: :	

Db 140 LELPYDWTHTLLEVPLODAREAGQURAGVDPAAAAARSLVAAPFGMQHVSNDLHORADIME 199

RESULT 8
ADH39748
ID ADH39748 standard; protein: 228 AA.

AC	ADH39748;
XX	
DT	11-MAR-2004 (first entry)

XX actinomycetes-originated plasmid; linear plasmid; pSLA2-L;
KW *Streptomyces rochei*; secondary metabolite; antibiotic; vitamin;
KW lankamycin; mithramycin-like substance; carotenoid lankacidin.
KW

OS Streptomyces rochei.

AA
PN
WO2004001039-A1.

31-DEC-2003.

PF 19-JUN-2003; 2003WO-JP007767.

20-JUN-2002; 2002JP-00179345.

XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

Kinashi H:

AA
DR
WPI; 2004-082503/08.

Actinomycetes-originated plasmids for producing secondary metabolites e.g. antibiotics and vitamins like lankamycins, lankacidins, mithramycin-like substances and carotenoids.

PS Claim 1; SEQ ID NO 79; 499bp; Japanese.

The present invention describes an actinomycetes-originated plasmid is a linear plasmid (pSLA2-L) originating in *Streptomyces rochei*, which has a base sequence capable of providing any of the protein sequences of SEQ ID NO:1 to 143 (ADH39670 to ADH39812), or one based on these sequences but with some amino acids deleted, substituted or added and capable of producing a secondary metabolite. Also described is a process for producing a secondary metabolite by using the plasmid. The plasmid can be used for producing secondary metabolites e.g. antibiotics and vitamins like lankamycins, mithramycin-like substances, carotenoids and lankadins. Such secondary metabolites are produced selectively and efficiently. The present sequence represents a *Streptomyces rochei* protein.

Sequence 228 AA;

Query Match	20.3%	Score 236;	DB 8;	Length 228;
Best Local Similarity	33.7%	Pred. NO. 2.6e-17;		
Matches	69;	Conservative	35;	Mismatches 85;
				Indels 16;
				Gaps 6;

[illegible]

Qy 69 LAPQTLLHVAVPESGSKA---QELVDLTMLVAHGMLHDPIIRAGTRIALDQGAVDFSDANP 125
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 64 ADAAVRLVPVHGAETGALQAVDYATYGLGSFLDDIVVRAGFELG-SQFALG-PEARM 121

Qy

126 FGWGDICALLAAQAQRGEVLPHVNPKKTGDFIVGCFTGIQAVSVTSDRQDLGRISV 185
| : | | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db

122 LHYWHHWVDOTLHRAKAEGOLAPGVSPGDAAKA VYATVGIRVLA-VGDPBWLAEERRIAA 180

Qy	:	186 MNHVLPSIVPASMLTWIETGEERI 210
Db	:	181 FWELLASIA-----AEPGRIRV 198

RESULT 9
ADH39815
ID ADH39815 standard; protein; 80 AA.

ADH39815;	
11-MAR-2004	(first entry)

actinomyces-originated plasmid; linear plasmid; pSLA2-L;
 Streptomyces rochei; secondary metabolite; antibiotic; vitamin;
 lankamycin; mithramycin-like substance; carotenoid; lankacidin.

OS Unidentified

PN WQ2004001039-A1.

PD 31-DEC-2003.

PF 19-JUN-2003; 2003WO-JP007767.

PR 20-JUN-2002; 2002JP-00179345.

PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

PI Kinashi H;

DR WPI; 2004-082503/08.

PT Actinomycetes-originated plasmids for producing secondary metabolites e.g.
PT antibiotics and vitamins like lankamycins, lankacidins, mithramycin-like
PT substances and carotenoids.

PS Disclosure; Fig 6A; 499pp; Japanese.

The present invention describes an actinomycetes-originated plasmid is a linear plasmid (pSUA2-L) originating in *Streptomyces rochei*, which has a base sequence capable of providing any of the protein sequences of SEQ ID NO:1 to 143 (ADH39670 to ADH39812), or one based on these sequences but with some amino acids deleted, substituted or added and capable of producing a secondary metabolite. Also described is a process for producing a secondary metabolite by using the plasmid. The plasmid can be used for producing secondary metabolites e.g. antibiotics and vitamins like lankamycins, mithramycin-like substances, carotenoids and lankadins. Such secondary metabolites are produced selectively and efficiently. The present sequence represents an amino acid sequence given in the exemplification of the present invention.

Sequence 80 AA;

Query Match	20.0%	Score 233;	DB 8;	Length 80;
Best Local Similarity	61.0%	Pred. No. 1.2e-17;		
Matches	47;	Conservative	14;	Mismatches 16;
				Indels 0;
				Gaps 0;

Qy 12 QERAVTRQAIIVRAAASVFDYGFEEATVAEILSRASVTKGAMYFHFASFKEELARGVLAE 71
 | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Dd 4 OVRAIRTRQAISAAARVFERGYOATISEITLVAGVTKGALYFHFOKEDLAOGVLT A 63

QY 72 QTLHVAVPESGSKAQEL 88
Db 64 QNEDLLLPERPAPKLOEV 80

RESULT 10
ADH39813
ID ADH39813 standard: protein: 80 AA.

XX ADH39813;
AC
DT 11-MAR-2004 (first entry)
DE
DE ArpA amino acid sequence.
XX
XX actinomycetes-originated plasmid; linear plasmid; pSLA2-L;
KW Streptomyces rochei; secondary metabolite; antibiotic; vitamin;
KW lankamycin; mithramycin-like substance; carotenoid; lankacidin.
XX
OS Unidentified.
XX
XX WO2004001039-A1.
XX
PD 31-DEC-2003.
XX
XX 19-JUN-2003; 2003WO-JP007767.
XX
XX 20-JUN-2002; 2002JP-00179345.
XX
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
XX Kinashi H;
XX
XX WPI; 2004-082503/08.
XX
XX Actinomycetes-originated plasmids for producing secondary metabolites e.g.
PT antibiotics and vitamins like lankamycins, lankacidins, mithramycin-like
PT substances and carotenoids.
XX
XX Disclosure; Fig 6A; 499pp; Japanese.
XX
XX The present invention describes an actinomycetes-originated plasmid is a
CC linear plasmid (pSLA2-L) originating in Streptomyces rochei, which has a
CC base sequence capable of providing any of the protein sequences of SEQ ID
CC NO:1 to 143 (ADH39670 to ADH39812), or one based on these sequences but
CC with some amino acids deleted, substituted or added and capable of
CC producing a secondary metabolite. Also described is a process for
CC producing a secondary metabolite by using the plasmid. The plasmid can be
CC used for producing secondary metabolites e.g. antibiotics and vitamins
CC like lankamycins, mithramycin-like substances, carotenoids and
CC lankacidins. Such secondary metabolites are produced selectively and
CC efficiently. The present sequence represents an amino acid sequence given
CC in the exemplification of the present invention.
XX
XX Sequence 80 AA;
SQ
Query Match 19.2%; Score 224; DB 8; Length 80;
Best Local Similarity 57.7%; Pred. No. 1.2e-16;
Matches 45; Conservative 14; Mismatches 19; Indels 0; Gaps 0;
QY 11 RERAVRTQAIIVRAAASVFDEYGFPAATVAEILSPASVTKGMYPHFASKBELARGVLA 70
DB 3 KQARAVQTVRSIVDAASVFDDYGYERAAISEILREAKVTKGALYHFFASKELIAQAIMD 62
QY 71 EQLTHVAVPESGSKAQEL 88
DB 63 EQISTVEFEQSGSPLQSL 80
RESULT 11
ADH39743
ID ADH39743 standard; protein; 213 AA.
XX
AC ADH39743;
XX
XX 11-MAR-2004 (first entry)
DT
XX Streptomyces rochei ORF74 protein SEQ ID NO:74.
DE
XX actinomycetes-originated plasmid; linear plasmid; pSLA2-L;
KW Streptomyces rochei; secondary metabolite; antibiotic; vitamin;
KW

lankamycin; mithramycin-like substance; carotenoid; lankacidin.
XX
XX Streptomyces rochei.
XX
XX WO2004001039-A1.
XX
XX 31-DEC-2003.
XX
XX 19-JUN-2003; 2003WO-JP007767.
XX
XX 20-JUN-2002; 2002JP-00179345.
XX
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
XX Kinashi H;
XX
XX WPI; 2004-082503/08.
XX
XX Actinomycetes-originated plasmids for producing secondary metabolites e.g.
PT antibiotics and vitamins like lankamycins, lankacidins, mithramycin-like
PT substances and carotenoids.
XX
XX Claim 1; SEQ ID NO 74; 499pp; Japanese.
XX
XX The present invention describes an actinomycetes-originated plasmid is a
CC linear plasmid (pSLA2-L) originating in Streptomyces rochei, which has a
CC base sequence capable of providing any of the protein sequences of SEQ ID
CC NO:1 to 143 (ADH39670 to ADH39812), or one based on these sequences but
CC with some amino acids deleted, substituted or added and capable of
CC producing a secondary metabolite. Also described is a process for
CC producing a secondary metabolite by using the plasmid. The plasmid can be
CC used for producing secondary metabolites e.g. antibiotics and vitamins
CC like lankamycins, mithramycin-like substances, carotenoids and
CC lankacidins. Such secondary metabolites are produced selectively and
CC efficiently. The present sequence represents a Streptomyces rochei
CC protein.
XX
XX Sequence 213 AA;
SQ
Query Match 18.4%; Score 214; DB 8; Length 213;
Best Local Similarity 29.1%; Pred. No. 6.5e-15;
Matches 65; Conservative 35; Mismatches 101; Indels 22; Gaps 4;
QY 12 QERAVRTQAIIVRAAASVFDEYGFPAATVAEILSPASVTKGMYPHFASKBELARGV--- 68
DB 2 QTRSERTRSLIQAGRGQFGRHGYAGATLDRIAADAGLTGKALYHFFASKELVLAFAVORE 61
QY 69 -----LAEQTLHVAVPESGSKAQELVDLTMLVAHGMHLDPILRAGTRLALDQGVDF 120
DB 62 GGARLASFLREREHEGAPP-----VQILDMTHLIRALREDPLVGAGVRIAQDRMDGP 116
QY 121 SDANPFGWGDICACOLLAAEQERGEVLPHPVNPCKTGDFIVGCTGLQAVSRVTSRQDQIG 180
DB 117 PATGLYRANDEHVELRLLRAGEAGALRAAGRGEAGALLWTTCGLAAVAGAL-PAGEDLD 175
QY 181 HRISVMNHLVPSIVPASMLTWIETGEERIGKVAIAAEAA 223
DB 176 RLTLALWDLTLVLVSQD-----EAARYRTHAPAPPGLAGAA 213
RESULT 12
AAU04039
ID AAU04039 standard; protein; 203 AA.
XX
XX AC AAU04039;
XX
XX 11-SEP-2003 (revised)
DT 23-OCT-2001 (first entry)
XX
XX Streptococcus coelicolor MmyR protein.
XX
XX SCPI; methylenomycin cluster; mmc; MmyR; MmfP; MmfH; MmfL; MmfR; MmyT;
KW MmyO; MmyG; MmyJ; Mmr; heterologous gene expression.
KW

```
XX OS Streptomyces coelicolor.
XX PN WO200148228-A1.
XX PD 05-JUL-2001.
XX PF 20-DEC-2000; 2000WO-GB004972.
XX PR 23-DEC-1999; 99GB-00030477.
XX PA (PLAN-) PLANT BIOSCIENCE LTD.
XX PI Chater KF, Bruton CJ, O'rouke SJ, Wietzorrek AW;
XX DR WPI; 2001-425675/45.
XX DR N-PSDB; AAS07627, AAU04045, AAU04046.
XX PT Novel expression cassette for expressing a nucleic acid of interest,
XX PT derived from the regulatory region of methylenomycin gene cluster of SCP1
XX PT plasmid of Streptomyces coelicolor A3(2).
XX PS Claim 25; Fig 8a; 142pp; English.
XX CC The sequence represents the MmyR protein encoded by the mmyR gene carried
XX CC on the expression cassette present on plasmid SCP1. The expression
XX CC cassette is the regulatory region of the methylenomycin cluster (mmc)
XX CC from Streptomyces coelicolor A3(2), which encodes the MmyR, MmfP, MmfH,
XX CC MmfL, MmfR, MmyT, MmyG, MmyJ and partial Mmr polypeptides. The
XX CC expression cassette is useful for expressing a nucleic acid of interest,
XX CC substantially only when the host cell culture reaches high cell density
XX CC at or close to the stationary phase of host cell culture. In particular
XX CC the system is useful in regulating methylenomycin production. Reduced or
XX CC no expression of the nucleic acid of interest is observed earlier in
XX CC growth, avoiding toxic effects of some gene products on growth and the
XX CC cluster naturally present on a highly transmissible plasmid permits
XX CC properly regulated expression in diverse Streptomyces host and the
XX CC expression is driven by a strong promoter, leading to high yield of the
XX CC desired end product. (Updated on 11-SEP-2003 to standardise OS field)
XX SQ Sequence 203 AA;

Query Match 13.9%; Score 161.5; DB 4; Length 203;
Best Local Similarity 34.1%; Pred. No. 4e-09;
Matches 44; Conservative 19; Mismatches 45; Indels 21; Gaps 3;

QY 10 VROERAVTRQAIIVRAAASVDEYGFEEATVAEILSRASVTKGMYPFHASKKEELARGVL 69
DQ 1 VQOARAVTRDQVLDAAAEFEALHGYAGTNLATVAVRTGTMGALYGHFPFKKALADELV 60
QY 70 AEQT-----LHVAVPSPSGSKAQLVDITLVAHGMHLDPILRAGTRIALD---- 114
DQ 61 SQSTETWNTIGRSIAETACAPETALRA-----LVLAVSRQMKHDIRFRAALRLAADCTMP 115
QY 115 -QGAVDFSD 122
DQ 116 AGGAPDLLD 124

RESULT 13
ID ADA34808
AC ADA34808 standard; protein; 211 AA.
XX ADA34808;
XX 20-NOV-2003 (first entry)
XX Acinetobacter baumannii protein #1969.
XX Acinetobacter baumannii; bacterial disease; antibiotic; vaccine;
XX plant biocontrol agent.
```

```
OS Acinetobacter baumannii.
XX PN US562958-B1.
XX PD 13-MAY-2003.
XX PF 04-JUN-1999; 99US-00328352.
XX PR 09-JUN-1998; 98US-0088701P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Breton G, Bush D;
XX DR WPI; 2003-576092/54.
XX DR N-PSDB; ADA30682.
XX PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
XX PT for diagnosing a bacterial disease, as components of antibacterial
XX PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
XX PT plants.
XX PS Example; SEQ ID NO 6095; 328pp; English.
XX CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
XX CC The A. baumannii nucleic acids and polypeptides are useful as reagents
XX CC for diagnosing a bacterial disease, as components of antibacterial
XX CC vaccines, as targets for antibacterial drugs, to detect the presence of
XX CC A. baumannii and other Acinetobacter species in a sample, in screening
XX CC compounds for the ability to interfere with the A. baumannii life cycle
XX CC or to inhibit A. baumannii infection, and as biocontrol agents for an A.
XX CC baumannii protein.
XX SQ Sequence 211 AA;

Query Match 11.9%; Score 138.5; DB 6; Length 211;
Best Local Similarity 22.9%; Pred. No. 1.5e-06;
Matches 49; Conservative 34; Mismatches 74; Indels 57; Gaps 5;

QY 1 MAVRHERVAVRQERAVRTRQAIIVRAAASVDEYGFEEATVAEILSRASVTKGMYPFHAFAS 60
DQ 4 LIVVYDECKRLNKSETTRQIHLDTSFELVHLKGFVGVLQELIKACDVPKGSFYHYPAS 63
QY 61 KEELARGVL-----AEQTLHVAV-----PESGSKAQE--LVDL 91
DQ 64 KEAFGALLLEQYMWYKVRMEQLWQHSEQSAHARLMALWQAWIDDPVHGSWAENCLIVKL 123
QY 92 TMLVA-----HGMHLDPILRAGTRIALDQGVDFSDANPFGEWGDICALLAEQERGE 145
DQ 124 AAEVSDLSSEDMRQILNDGVHKLTORLAL-----LUKEGQOEGS 161
QY 146 VLPHVNPVKTGDFIVGCTFGLQAVSRVTSDDQLD 179
DQ 162 IPKHIDPLKTAQVMYQLWLGAAALLKLSQDKAHL 195

RESULT 14
AAW42428
ID AAW42428 standard; protein; 195 AA.
XX AAW42428;
XX 06-JUL-1998 (first entry)
XX Glucuronide repressor GusR.
XX Glucuronide repressor; GusR; transgene detection; assay.
XX Escherichia coli.
XX Key Location/Qualifiers
XX Domain 1..63
```

FT Domain /note= "operator binding domain"
FT 64. 154
FT /note= "glucuronide binding domain (may extend to residue
FT 195)"
PN WO9749813-A2.
PD 31-DEC-1997.
XX 26-JUN-1997; 97WO-US011427.
XX 26-JUN-1996; 96US-0020621P.
XX (CAMB-) CAMBIA BIOSYSTEMS LLC.
XX Jefferson RA, Wilson KJ, Leader M;
PI WPI; 1998-077176/07.
DR N-PSDB; AAV03424, AAV03430.
XX Glucuronide repressor gene - useful to develop products to detect
PT glucuronide(s), control gene expression or identify glucuronide transport
PT proteins.
XX Claim 12; Page 40; 84pp; English.

XX This polypeptide nucleic acid molecule comprises the glucuronide
CC repressor (GusR) of Escherichia coli. The existence of the GusR gene (see
CC AAV03424) was established by genetic and biochemical experiments and
CC genetically mapped to a region upstream of the glucuronidase gene gusA
CC (see also AAV03430). GusR acts by binding to gusA operator sequences (see
CC AAV03429), thus preventing transcription, this repression being relieved
CC when a glucuronide substrate binds to the repressor and inactivates it.
CC Overexpression of the GusR gene product is achieved by cloning the coding
CC region in an expression vector. Vectors comprising a nucleic acid
CC molecule encoding GusR, its glucuronide binding domain or a fusion
CC protein, especially a binary Agrobacterium tumefaciens plasmid vector, as
CC well as plant, animal, fungal and bacterial host cell transformed by such
CC a vector are claimed. The repressor is used in claimed methods for
CC controlling expression of a transgene, detecting glucuronides in a sample
CC and isolating glucuronides from a sample

XX Sequence 195 AA;

Query Match 11.9%; Score 138; DB 2; Length 195;
Best Local Similarity 26.3%; Pred. No. 1.5e-06;
Matches 51; Conservative 31; Mismatches 72; Indels 40; Gaps 8;
QY 12 QERAVTRQAIIVRAAASVFDEYGFEEATVAEILSRASVTGKAMYFHFASKEELARG-VLA 70
Db 5 QTEAQPTTRILNAREIFESENGFHSASMKACKSCAISPGLYHHFISKEALIQAILQ 64
QY 71 EOTLHVA---VPESG-----SKAQELVDLT-----MLVAHGMLHDPILRAGTRL 111
Db 65 DQERALARFREPTEGHIHFVDYMWESIVSLTHEAFQGBALVVEIMAEQM-RNQVAAMLK- 122
QY 112 ALDQGAVFDSANPFGWGDICQLAAEQERGEVLPVHNPVKTG-----DFIVGCTGLQ 167
Db 123 -----NKHWTITEFVAQRMRDAQKGEISPDINTAMTSRLLLDLTYGVLAIDIE 170
QY 168 A--VSRVTSRQDL 179
Db 171 AEDLAREASPAQGL 184

RESULT 15
ABU16923
ID ABU16923 standard; protein; 200 AA.

XX ABU16923;
XX 19-JUN-2003 (first entry)
XX

DE Protein encoded by Prokaryotic essential gene #2450.
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Acinetobacter baumannii.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX N-PSDB; ACA20793.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 44847; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 200 AA;

Query Match 11.6%; Score 135.5; DB 6; Length 200;
Best Local Similarity 23.5%; Pred. No. 3e-06;
Matches 48; Conservative 30; Mismatches 69; Indels 57; Gaps 5;
QY 11 QERAVTRQAIIVRAAASVFDEYGFEEATVAEILSRASVTGKAMYFHFASKEELARGVL- 69
Db 3 RLNKSETTRQHLDTSLFELVHLKGFVGLQELKACDVPKGSFYHFYFASKEAFGCALLE 62

```
Qy 70 -----ABQTLHVAV-----PESGSKAQE--LVDLTMLVA----- 96
Db 63 QYMANYKVRMEQLWOHSEQSAHARLMAWQAWIDDPVHGSAWENCLIVKLAAEVSDLSED 122
Qy 97 -HGMHLDPILRACGTRIALDOGAVDFSDANPFGEWGDICACQLLAEAOERGEVLPHVNPCKT 155
Db 123 MRQILNDGVHKLRTORLAL-----LLKEGQOEGSIPKHIEPLKT 160
Qy 156 GDFIVGCFGLQAVSRVTSRQDL 179
Db 161 AQVMYQLWLGAALLTKLSQDKAHL 184
```

Search completed: March 6, 2005, 22:29:13
Job time : 152 secs

This Page Blank (uspo)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2005, 22:16:17 ; Search time 38 Seconds
(without alignments)
455.752 Million cell updates/sec

Title: US-10-049-710A-2
Perfect score: 1164
Sequence: 1 MAVRHERVAVRGERAVTRQ.....VAAAEAAAEASAEASDE 232

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A-COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5S-COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A-COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B-COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCUTUS-COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138.5	11.9	211	4	US-09-328-352-6095 Sequence 6095, Ap
2	138	11.9	195	2	US-08-882-704A-2 Sequence 2, Appli
3	138	11.9	195	4	US-09-151-957-2 Sequence 2, Appli
4	134	11.5	201	4	US-09-252-991A-30890 Sequence 30890, A
5	133	11.4	190	3	US-09-303-120B-6 Sequence 6, Appli
6	133	11.4	190	3	US-09-820-576-6 Sequence 6, Appli
7	133	11.4	190	4	US-09-966-608-6 Sequence 6, Appli
8	121	10.4	192	3	US-09-303-120B-8 Sequence 8, Appli
9	121	10.4	192	3	US-09-820-576-8 Sequence 8, Appli
10	121	10.4	192	4	US-09-966-608-8 Sequence 8, Appli
11	119	10.2	229	4	US-09-489-039A-10340 Sequence 10340, A
12	118.5	10.2	216	4	US-09-252-991A-20124 Sequence 20124, A
13	115.5	9.9	230	4	US-09-489-039A-13695 Sequence 13695, A
14	115.5	9.9	234	4	US-09-543-681A-4712 Sequence 4712, Ap
15	115	9.9	254	4	US-09-252-991A-18058 Sequence 18058, A
16	114	9.8	230	4	US-09-252-991A-22064 Sequence 22064, A
17	113.5	9.8	213	4	US-09-252-991A-28039 Sequence 28039, A
18	113	9.7	210	3	US-09-199-637A-1172 Sequence 172, App
19	113	9.7	243	4	US-09-252-991A-26765 Sequence 26765, A
20	112.5	9.7	235	4	US-09-489-039A-14058 Sequence 14058, A
21	109.5	9.4	207	4	US-08-328-352-5603 Sequence 5603, Ap
22	109	9.4	202	4	US-09-502-540-14870 Sequence 14870, A
23	108.5	9.3	198	4	US-09-976-451-5 Sequence 5, Appli
24	108.5	9.3	223	4	US-09-543-681A-6126 Sequence 6126, Ap
25	108	9.3	204	4	US-09-134-000C-4883 Sequence 4883, Ap
26	108	9.3	206	4	US-09-328-352-6491 Sequence 6491, Ap
27	108	9.3	543	4	US-09-252-991A-27586 Sequence 27586, A

28	107.5	9.2	400	4	US-09-252-991A-32313 Sequence 32313, A
29	106	9.1	195	3	US-09-134-001C-3548 Sequence 3548, Ap
30	105	9.0	200	3	US-09-303-120B-7 Sequence 7, Appli
31	105	9.0	200	3	US-09-820-576-7 Sequence 7, Appli
32	105	9.0	200	4	US-09-966-608-7 Sequence 7, Appli
33	105	9.0	205	4	US-09-252-991A-25485 Sequence 25485, A
34	104	8.9	205	4	US-09-328-352-4800 Sequence 4800, Ap
35	104	8.9	217	4	US-09-489-039A-13254 Sequence 13254, A
36	104	8.9	245	4	US-09-252-991A-26292 Sequence 26292, A
37	102	8.8	199	4	US-09-543-681A-4892 Sequence 4892, Ap
38	102	8.8	211	4	US-09-107-532A-3886 Sequence 3886, Ap
39	101.5	8.7	204	4	US-09-252-991A-30118 Sequence 30118, A
40	101	8.7	189	4	US-09-704-725-10 Sequence 10, Appl
41	101	8.7	243	4	US-09-902-540-12333 Sequence 12333, A
42	100.5	8.6	207	4	US-09-489-039A-10404 Sequence 10404, A
43	100.5	8.6	280	4	US-09-252-991A-30349 Sequence 30349, A
44	100.5	8.6	312	4	US-09-252-991A-20753 Sequence 20753, A
45	100	8.6	195	4	US-09-252-991A-26972 Sequence 26972, A

ALIGNMENTS

RESULT 1
US-09-328-352-6095
; Sequence 6095, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6095
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6095

Query Match 11.9%; Score 138.5; DB 4; Length 211;
Best Local Similarity 22.9%; Pred. No. 4e-08;
Matches 49; Conservative 34; Mismatches 74; Indels 57; Gaps 5;
QY 1 MAVRHERVAVRGERAVTRQAIIVRAAASVFDEYGFBAATVAEILSRASVTKGAMYHFHAS 60
DB 4 LIVVYDECMKRLNKSETTRQHILDTSFELVHKGFVGVGLQEILKACDVPKGSFYHFAS 63
QY 61 KEELARGVL-----AQTLHVAV-----PESGSKAOE--LVDL 91
DB 64 KEAFGCALLEQYMANYSKVRMEQLWQHSQSAHRLMALWQAWIDDPVHGSWAENCLIVKL 123
QY 92 TMLVA-----HGLMDHPILRAGTRIALDQGVDFSDANPFGEWGDICALLAEAOERGE 145
DB 124 AAEVSDLSMDRQILNDGVHKLQRLAL-----LLKEGQOEGS 161
QY 146 VLPHVNPKITGTFIVGCTFTGLQAVSRVTSRDQDL 179
DB 162 IPKHIDPLTAQVMYQLWLGAALLTKLSQDKAHL 195

RESULT 2
US-08-882-704A-2
; Sequence 2, Application US/08882704A
; Patent No. 5879906
; GENERAL INFORMATION:
; APPLICANT: Jefferson, Richard A.
; APPLICANT: Wilson, Katherine J.
; APPLICANT: Leader, Michael
; TITLE OF INVENTION: GLUCURONIDE REPRESSORS AND USES THEREOF
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:

ADDRESSES: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/882,704A
FILING DATE: 25-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5879906tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 190106.404
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 195 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear

US-08-882-704A-2

Query Match 11.9%; Score 138; DB 2; Length 195;
Best Local Similarity 26.3%; Pred. No. 4.1e-08;
Matches 51; Conservative 31; Mismatches 72; Indels 40; Gaps 8;

QY 12 QERAVRTRQAIIVRAAASVFDEYGFEEATVAEILSRASVTKGAMYPHFASKEELARG-VLA 70
Db 5 QTEAQTTRILNAREIFSENGFHSAMKAICKSAISPGTLYHHFISKEALIQAILIQ 64

QY 71 EQTLHVA---VPESG-----SKAQLVDLT-----MLVAHGMHDPILRAGTRL 111
Db 65 DQERALARFREPIEGIHFDYVNVESIVSLTHEAFQGRALVVEIMAEGM-RNQVAAAMLK- 122

QY 112 ALDQGAVDSDANPFGWGDICALLAEAOERGEVLPHVNPCKTG----DFIVGCFGTGLQ 167
Db 123 -----NKMTITFEVQMRDAQKQGEISPDINTAMTSRLLLDLTYGVLADE 170

QY 168 A--VSRVTSRDQDL 179
Db 171 AEDLAREASFAQGL 184

RESULT 3
US-09-151-957-2
Sequence 2, Application US/09151957
Patent No. 6429292
GENERAL INFORMATION:
APPLICANT: Jefferson, Richard A.
Leader, Michael
Willson, Katherine J.
TITLE OF INVENTION: GLUCURONIDE REPRESSORS AND USES THEREOF
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/151,957
FILING DATE: 11-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/882,704
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: No. 6429292tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 190106.404
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 195 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-151-957-2

Query Match 11.9%; Score 138; DB 4; Length 195;
Best Local Similarity 26.3%; Pred. No. 4.1e-08;
Matches 51; Conservative 31; Mismatches 72; Indels 40; Gaps 8;

QY 12 QERAVRTRQAIIVRAAASVFDEYGFEEATVAEILSRASVTKGAMYPHFASKEELARG-VLA 70
Db 5 QTEAQTTRILNAREIFSENGFHSAMKAICKSAISPGTLYHHFISKEALIQAILIQ 64

QY 71 EQTLHVA---VPESG-----SKAQLVDLT-----MLVAHGMHDPILRAGTRL 111
Db 65 DQERALARFREPIEGIHFDYVNVESIVSLTHEAFQGRALVVEIMAEGM-RNQVAAAMLK- 122

QY 112 ALDQGAVDSDANPFGWGDICALLAEAOERGEVLPHVNPCKTG----DFIVGCFGTGLQ 167
Db 123 -----NKMTITFEVQMRDAQKQGEISPDINTAMTSRLLLDLTYGVLADE 170

QY 168 A--VSRVTSRDQDL 179
Db 171 AEDLAREASFAQGL 184

RESULT 4
US-09-252-991A-30890
Sequence 30890, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 30890
LENGTH: 201
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30890

Query Match 11.5%; Score 134; DB 4; Length 201;
Best Local Similarity 34.7%; Pred. No. 1.3e-07;
Matches 33; Conservative 15; Mismatches 35; Indels 12; Gaps 2;

QY 11 QERAVRTRQAIIVRAAASVFDEYGFEEATVAEILSRASVTKGAMYPHFASKEELARGVLA 70
Db 11 QQENAEATREALLESALSALFIEHGYGVGSIDAIAREARVTKGAFYHHFGSKQELLAECEY 70

```
QY 71 EQTLHVA-----VP-----ESGSKAQLVLDLTM 93
Db 71 RQVRTIAEDLDVRPAHVWKWAEAAALAEAFIDSV 105

RESULT 5
US-09-303-120B-6
; Sequence 6, Application US/09303120B
; Patent No. 6309643
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; TITLE OF INVENTION: IBID-Associated Microbial Antigens and Methods of Using
; FILE OF INVENTION: Same
; CURRENT APPLICATION NUMBER: US/09/303,120B
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Clostridium pasteurianum
US-09-303-120B-6

Query Match 11.4%; Score 133; DB 3; Length 190;
Best Local Similarity 24.3%; Pred. No. 1.6e-07;
Matches 49; Conservative 33; Mismatches 78; Indels 42; Gaps 7;

QY 17 RTRQAIVRAAASVFDEYGFEEATVAAILSRASVTYKGMATYFHFASKELARGVLAEE--QTL 74
Db 3 KTKDNIFYSAIKVFSNNGYNGATMDIASNAGVAKGTLYYHFKSKEEIKFYIIIEGVNLM 62

QY 75 HVAVPESGSKAQLVLDLTMVAHGMHLDPILRAGTTLALDQGVAVDFSDANPFGEWGDICA 134
Db 63 KNEIDEATDKERTALEKLVACR-----VOLNLIYKNRDFPKVIASQLWKGELR 111

QY 135 QL-----LAAEQERGEVLPHVNPVKTKGD--FIVGCTFG-LQAVSR---V 172
Db 112 QLELRDIRMNVYVHIEEFVKDAMEAGSI-----KKGNSLFVAYAFGLTCLCSVSLYEVI 164

QY 173 TSDRQDLGHRISVMNHNVLPSI 194
Db 165 NAENDNINNTIENLMNYILNGI 186

US-09-820-576-6
; Sequence 6, Application US/09820576
; Patent No. 6320037
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; TITLE OF INVENTION: IBID-Associated Microbial Nucleic Acid
; FILE OF INVENTION: Molecules
; FILE REFERENCE: P-PM 4646
; CURRENT APPLICATION NUMBER: US/09/820,576
; PRIOR FILING DATE: 2001-03-28
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Clostridium pasteurianum
US-09-820-576-6

Query Match 11.4%; Score 133; DB 3; Length 190;
Best Local Similarity 24.3%; Pred. No. 1.6e-07;
Matches 49; Conservative 33; Mismatches 78; Indels 42; Gaps 7;

QY 17 RTRQAIVRAAASVFDEYGFEEATVAAILSRASVTYKGMATYFHFASKELARGVLAEE--QTL 74
Db 3 KTKDNIFYSAIKVFSNNGYNGATMDIASNAGVAKGTLYYHFKSKEEIKFYIIIEGVNLM 62

QY 75 HVAVPESGSKAQLVLDLTMVAHGMHLDPILRAGTTLALDQGVAVDFSDANPFGEWGDICA 134
Db 63 KNEIDEATDKERTALEKLVACR-----VOLNLIYKNRDFPKVIASQLWKGELR 111

QY 135 QL-----LAAEQERGEVLPHVNPVKTKGD--FIVGCTFG-LQAVSR---V 172
Db 112 QLELRDIRMNVYVHIEEFVKDAMEAGSI-----KKGNSLFVAYAFGLTCLCSVSLYEVI 164

QY 173 TSDRQDLGHRISVMNHNVLPSI 194
Db 165 NAENDNINNTIENLMNYILNGI 186

US-09-820-576-6
; Sequence 6, Application US/09820576
; Patent No. 6320037
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; TITLE OF INVENTION: IBID-Associated Microbial Nucleic Acid
; FILE OF INVENTION: Molecules
; FILE REFERENCE: P-PM 4646
; CURRENT APPLICATION NUMBER: US/09/820,576
; PRIOR FILING DATE: 2001-03-28
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Clostridium pasteurianum
US-09-820-576-6

Query Match 11.4%; Score 133; DB 3; Length 190;
Best Local Similarity 24.3%; Pred. No. 1.6e-07;
Matches 49; Conservative 33; Mismatches 78; Indels 42; Gaps 7;
```

```
QY 17 RTRQAIVRAAASVFDEYGFEEATVAAILSRASVTYKGMATYFHFASKELARGVLAEE--QTL 74
Db 3 KTKDNIFYSAIKVFSNNGYNGATMDIASNAGVAKGTLYYHFKSKEEIKFYIIIEGVNLM 62

QY 75 HVAVPESGSKAQLVLDLTMVAHGMHLDPILRAGTTLALDQGVAVDFSDANPFGEWGDICA 134
Db 63 KNEIDEATDKERTALEKLVACR-----VOLNLIYKNRDFPKVIASQLWKGELR 111

QY 135 QL-----LAAEQERGEVLPHVNPVKTKGD--FIVGCTFG-LQAVSR---V 172
Db 112 QLELRDIRMNVYVHIEEFVKDAMEAGSI-----KKGNSLFVAYAFGLTCLCSVSLYEVI 164

QY 173 TSDRQDLGHRISVMNHNVLPSI 194
Db 165 NAENDNINNTIENLMNYILNGI 186

US-09-966-608-6
; Sequence 6, Application US/09966608
; Patent No. 6759530
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; TITLE OF INVENTION: IBID-Associated Microbial Nucleic Acid
; FILE OF INVENTION: Molecules
; FILE REFERENCE: P-PM 4966
; CURRENT APPLICATION NUMBER: US/09/966,608
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 09/303,120
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 09/820,576
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Clostridium pasteurianum
US-09-966-608-6

Query Match 11.4%; Score 133; DB 4; Length 190;
Best Local Similarity 24.3%; Pred. No. 1.6e-07;
Matches 49; Conservative 33; Mismatches 78; Indels 42; Gaps 7;

QY 17 RTRQAIVRAAASVFDEYGFEEATVAAILSRASVTYKGMATYFHFASKELARGVLAEE--QTL 74
Db 3 KTKDNIFYSAIKVFSNNGYNGATMDIASNAGVAKGTLYYHFKSKEEIKFYIIIEGVNLM 62

QY 75 HVAVPESGSKAQLVLDLTMVAHGMHLDPILRAGTTLALDQGVAVDFSDANPFGEWGDICA 134
Db 63 KNEIDEATDKERTALEKLVACR-----VOLNLIYKNRDFPKVIASQLWKGELR 111

QY 135 QL-----LAAEQERGEVLPHVNPVKTKGD--FIVGCTFG-LQAVSR---V 172
Db 112 QLELRDIRMNVYVHIEEFVKDAMEAGSI-----KKGNSLFVAYAFGLTCLCSVSLYEVI 164

QY 173 TSDRQDLGHRISVMNHNVLPSI 194
Db 165 NAENDNINNTIENLMNYILNGI 186

RESULT 8
US-09-303-120B-8
; Sequence 8, Application US/09303120B
; Patent No. 6309643
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; TITLE OF INVENTION: IBID-Associated Microbial Antigens and Methods of Using
; FILE OF INVENTION: Same
; FILE REFERENCE: P-PM 3478
; CURRENT APPLICATION NUMBER: US/09/303,120B
```

```
; TITLE OF INVENTION: Molecules
; FILE REFERENCE: P-PW 4966
; CURRENT APPLICATION NUMBER: US/09/966,608
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 09/303,120
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 09/820,576
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 192
; TYPE: PRF
; ORGANISM: Aulfex aeolicus
US-09-966-608-8

Query Match      10.4%; Score 121; DB 4; Length 192;
Best Local Similarity 24.7%; Pred.No. 4.3e-06;
Matches 41; Conservative 31; Mismatches 46; Indels 48; Gaps 7;

QY    13 ERAVTRQAIVRAAASVFDEYGEFAATVAEILSRASVTKGAMYTHFPASKEELARGVLAEQ 72
Db     ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY    9 EKRSDTKEKILLSALKLFSKGKETIMDKIAKEVGITGEGAIYRHFHTSKEEIIKLLLESI 68
Db     ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY    73 T-----LHVAVPESGSKAQELVDLTMLVAHGLMHPILRAGTRLRALDQGAVDFSDANP- 125
Db     ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY    69 TKELRHKLVEAL-QRGETDEELE-----SIVDTLIIDYAFSNPE 106
Db     ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY    126 ---FGEWGMDICALLAAEAQERGEV--LPHVNPKKTGDFIVGCFTGL 166
Db     ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY    107 SFRF-----LNLYHLKEYGEVKNL-----GELILKFLNGL 138
Db     ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

RESULT 11
US-09-489-039A-10340
; Sequence 10340, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10340
; LENGTH: 229
; TYPE: PRF
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10340

Query Match      10.2%; Score 119; DB 4; Length 229;
Best Local Similarity 26.8%; Pred.No. 9.8e-06;
Matches 50; Conservative 30; Mismatches 84; Indels 28; Gaps 7;

QY    3 VRHVRVAVQERAVRT-----RQAIIVRAAASVFDEYGEFAATVAEILSRASVT 50
Db     ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY    8 INYERVACRLSAMSETPRQGRPKDPPLTKQAILQAARKLFLEQGLE-VTTAEIARVAGVA 66
Db     ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY    51 KGAMYTHFPASKEELARGVLAEQTILHAVVPESGSKAQELVDLTMLVAHG----MLHDPIL 105
Db     ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY    67 KATLYANFSDKEHLIEAVLRQESDLTISDHDFQQRHLPDLEIVLATGYRFVRFINOREL 126
Db     ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY    106 RAGTRLALDQGVDFSD-----ANPFGEWGDTCAQLLAAEQERGEVLPHVNPKKTGDF 158
Db     ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY    127 TGWDRL-IASAAVRHPDLPGRFYAAGF-GRAQQMLEAIIAEIAG-TLRACDPOEAADE 183
Db     ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY    159 IVCGFTGLQAVS 170
Db     ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY    184 LAGLWLGMTSLA 195
Db     ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
```

RESULT 12
US-09-252-991A-20124
; Sequence 20124, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20124
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20124

Query Match 10.2%; Score 118.5; DB 4; Length 216;
Best Local Similarity 24.5%; Pred. No. 1e-05; Indels 53; Gaps 8;
Matches 47; Conservative 25; Mismatches 67

QY 12 QERAVR---TROAIVRAAASVFDEYCFEAAATVAEILSRASVTGKAMYFHFASKEELARG 67
DB 27 QEPSVRLKISRDQLQRCAGTFRYGYHCTTMDLSSACGLTKASFYHHYPNKEALLRD 86
QY 68 VLAEOQLHVAVPESGSKAQELVDLTMLVAHGMHLDPI-----RAGTRLALDQG- 116
DB 87 VL---EWTHT-----QRLAETLFSIA-----YDPLLTFRERLEKLGKRAARLFQDDSI 130
QY 117 -----AVDFSDANP-----FGWGDICQAQLAQAERGEVLPVHNPCKTGFIV 160
DB 131 GCLMGVVAVDASVGRSELMAPIRSLDDWAQAQLYRPAFDEAQAAL-----ERGRQLV 184
QY 161 GCFTGLQAVSRV 172
DB 185 ADEGAILLARI 196

RESULT 13
US-09-489-039A-13695
; Sequence 13695, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13695
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13695

Query Match 9.9%; Score 115.5; DB 4; Length 230;
Best Local Similarity 30.0%; Pred. No. 2.6e-05; Indels 27; Gaps 5;
Matches 36; Conservative 20; Mismatches 37

QY 12 QERAVRTRQAIIVRAAASVFDEYCFEAAATVAEILSRASVTGKAMYFHFASKEELARGVLAE 71
DB 20 KEAQRTRQLLIESAICQFALRGVNTTLTDIADAAGVTRGAVYWHFASKTELFENWQQ 79
QY 72 QTLHVAVPESGSKAQELVDLTMLVAHGMHLDPI--LR-----ACTRLALDQGVDFSDANP 125

DB 80 Q-----PPLRDLIQPSQAIEYE--HEPLNALRERFIAGLR-----YTAANP 118

RESULT 14
US-09-543-681A-4712
; Sequence 4712, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABI
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4712
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4712

Query Match 9.9%; Score 115.5; DB 4; Length 234;
Best Local Similarity 30.3%; Pred. No. 2.7e-05; Indels 13; Gaps 3;
Matches 40; Conservative 17; Mismatches 62

QY 11 ROERAVRTRQAIIVRAAASVFDEYCFEAAATVAEILSRASVTGKAMYFHFASKEELARGVLA 70
DB 34 QENSEQTRLALLEAQYLFFVNCYYDVSDISIRYARVTGKAFYHFFSNKALLRECYL 93
QY 71 EQLTHVA-----VPESGSKAQELVDLTMLVAHGMHLDPILRAGTRLALDQGVDFSDANP 125
DB 94 LQVGHAVQKLDDEVPTVDDKQELTALFSLCV-----DHIVQCKNELIPLQEVISVLG--- 145
QY 126 FGWGDICQAQL 137
DB 146 WKWEEDIDANIL 157

RESULT 15
US-09-252-991A-18058
; Sequence 18058, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18058
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18058

Query Match 9.9%; Score 115; DB 4; Length 254;
Best Local Similarity 25.6%; Pred. No. 3.5e-05; Indels 10; Gaps 4;
Matches 40; Conservative 30; Mismatches 76

QY 13 ERAVRTRQAIIVRAAASVFDEYCFEAAATVAEILSRASVTGKAMYFHFASKEELARGVLAEQ 72
DB 51 ERGRQRRRLDAATAQFLHGFEGTTLDMVIERAGSGRGTLYSSFGKEGLFAAVIAHM 110
QY 73 TRLHVAVPESGSKAQELVDLTMLVAH-----GMLHDPILRAGTRLALDQGVDFSDANP 126

Db 111 -IEEIFDDSDQPRPAATLSATLEHFGRRFLTSLDPRCQSLYRLVVAESPRFPAICKSF 169

Qy 127 GEWGDIQAQLLAQAQGEVLPVHNPXKKTGDFIVGC 162

Db 170 YEQGPQSYLL-SERLAAPVHMDDETL--YAVAC 202

Search completed: March 6, 2005, 22:35:27
Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2005, 22:27:08 ; Search time 128 Seconds
(without alignments)
596.394 Million cell updates/sec

Title: US-10-049-710A-2

Perfect score: 1164

Sequence: 1 MAVRHVAVRQERAVRTQ.....VAAAEAAAEAEASEASDE 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1391452 seqs, 329044822 residues

Total number of hits satisfying chosen parameters: 1391452

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTU5_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/2/pubpaa/US11_PUBCOMB.pep.*
20:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1164	100.0	232	14	US-10-049-710A-2
2	1159	99.6	232	15	US-10-168-663-9
3	495	42.5	221	15	US-10-168-663-7
4	476	40.9	215	14	US-10-017-471A-16
5	472	40.5	215	15	US-10-168-663-8
6	470.5	40.4	276	15	US-10-168-663-10
7	439.5	37.8	234	14	US-10-156-761-11242
8	272	23.4	215	15	US-10-168-663-5
9	257	22.1	215	15	US-10-168-663-6
10	248.5	21.3	218	14	US-10-156-761-11239
11	243	20.9	214	15	US-10-168-663-11
12	224	19.2	196	15	US-10-168-663-4
13	217.5	18.7	216	15	US-10-168-663-3

14	215.5	18.5	218	14	US-10-156-761-9807	Sequence 9807, Ap
15	207.5	17.8	287	14	US-10-156-761-11240	Sequence 11240, A
16	171.5	14.7	206	14	US-10-156-761-9805	Sequence 9805, Ap
17	161.5	13.9	203	15	US-10-168-663-12	Sequence 12, Appl
18	138	11.9	195	14	US-10-195-518-2	Sequence 2, Appl
19	138	11.9	227	14	US-10-156-761-13631	Sequence 13631, A
20	135.5	11.6	200	15	US-10-282-122A-44847	Sequence 44847, A
21	133	11.4	190	9	US-09-820-576-6	Sequence 6, Appl
22	133	11.4	190	9	US-09-966-608-6	Sequence 6, Appl
23	133	11.4	190	16	US-10-835-914-6	Sequence 6, Appl
24	132.5	11.4	295	15	US-10-282-122A-49782	Sequence 49782, A
25	131.5	11.3	273	15	US-10-282-122A-48142	Sequence 48142, A
26	131	11.3	199	14	US-10-156-761-11396	Sequence 11396, A
27	129.5	11.1	211	14	US-10-156-761-8108	Sequence 8108, Ap
28	127	10.9	197	14	US-10-156-761-12316	Sequence 12316, A
29	125	10.7	111	14	US-10-238-075-1557	Sequence 1557, Ap
30	124.5	10.7	253	15	US-10-461-194-128	Sequence 128, App
31	122.5	10.5	214	14	US-10-156-761-1977	Sequence 7977, Ap
32	121	10.4	192	9	US-09-820-576-8	Sequence 8, Appl
33	121	10.4	192	9	US-09-966-608-8	Sequence 8, Appl
34	121	10.4	192	16	US-10-835-914-8	Sequence 8, Appl
35	120.5	10.4	188	9	US-09-738-626-5200	Sequence 5200, Ap
36	120.5	10.4	188	17	US-10-495-066-36	Sequence 36, Appl
37	120	10.3	205	15	US-10-282-122A-49248	Sequence 49248, A
38	119.5	10.3	219	14	US-10-156-761-13447	Sequence 13447, A
39	119	10.2	194	14	US-10-156-761-8051	Sequence 8051, Ap
40	118.5	10.2	81	15	US-10-282-122A-50775	Sequence 50775, A
41	117.5	10.1	211	14	US-10-156-761-8127	Sequence 8127, Ap
42	117	10.1	220	14	US-10-156-761-10827	Sequence 10827, A
43	117	10.1	221	14	US-10-156-761-9317	Sequence 9317, Ap
44	115.5	9.9	189	14	US-10-238-075-1037	Sequence 1037, Ap
45	115.5	9.9	191	14	US-10-156-761-12235	Sequence 12235, A

ALIGNMENTS

RESULT 1

US-10-049-710A-2
; Sequence 2, Application US/10049710A
; Publication No. US20030126648A1
; GENERAL INFORMATION:
; APPLICANT: Shimyo, Atsuhiko
; APPLICANT: Kato, Kou
; APPLICANT: Yamada, Yasuhiro
; APPLICANT: Nihira, Takuya
; APPLICANT: Shindo, Takuya
; TITLE OF INVENTION: METHOD FOR INDUCTION OF GENE EXPRESSION IN PLANT AND PLANT TREAT
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 5405/18
; CURRENT APPLICATION NUMBER: US/10/049,710A
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: PCT/JP01/05096
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: JP 2000-180466
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Streptomyces virginiae
US-10-049-710A-2

Query Match 100.0%; Score 1164; DB 14; Length 232;

Best Local Similarity 100.0%; Pred. No. 5.9e-113; Indels 0; Gaps 0; Matches 232; Conservative 0; Mismatches 0;

QY 1 MAVRHVAVRQERAVRTQAIIVRAAASVDFYGFEEATVAEILSRASVTKGMYHFAS 60

Db 1 MAVRHVAVRQERAVRTQAIIVRAAASVDFYGFEEATVAEILSRASVTKGMYHFAS 60

QY 61 KEELARGVLAEGTLHVAVPESGSKAQELVDLTMLVAHGMHLDPILRAGTRRLDQGVDF 120

61	KEELARGVLAEGTTLHVA	PPESGSKAQLVDVLTMLVAHGMUHDPI	LLRAGTRFLADQGVADF	120	
121	SDANPFGWGWDICALLAE	QERGEREVLPHVNP	PKTGDPIVGCCTG	QAVSRVTSRDRQDLG	180
121	SDANPFGWGWDICALLAE	QERGEREVLPHVNP	PKTGDPIVGCCTG	QAVSRVTSRDRQDLG	180
181	HRISVYNNHVLPSI	VPASMLTWITGTGER	IGKVA	AAAAAEEAAASEAASDE	232
181	HRISVYNNHVLPSI	VPASMLTWITGTGER	IGKVA	AAAAAEEAAASEAASDE	232

```

RESULT 2
US-10-168-663-9
; Sequence 9, Application US/10168663
; Publication No. US20040086962A1
; GENERAL INFORMATION:
; APPLICANT: Plant Bioscience Limited
; APPLICANT: Chater, Keith F
; APPLICANT: Bruton, Celia J
; APPLICANT: O'Rourke, Sean J
; APPLICANT: Wietzorrek, Andreas W
; TITLE OF INVENTION: Methods and Materials Relating to Gene Expression
; FILE REFERENCE: 0380-P029090US
; CURRENT APPLICATION NUMBER: US/10/168,663
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: PCT/GB00/04972
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: GB 9930477.6
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Streptomyces sp.
US-10-168-663-9

```

RESULT 3
US-10-168-663-7
; Sequence 7, Application US/10168663
; Publication No. US20040086962A1
; GENERAL INFORMATION:
; APPLICANT: Plant Bioscience Limited
; APPLICANT: Chater, Keith F
; APPLICANT: Bruton, Selia J
; APPLICANT: O'Rourke, Sean J
; APPLICANT: Wietzorrek, Andreas W
; TITLE OF INVENTION: Methods and Materials Relating to Gene Expression
; FILE REFERENCE: 0380-P02909USO
; CURRENT APPLICATION NUMBER: US/10/168,663
; CURRENT FILING DATE: 2002-10-25

```

; PRIOR APPLICATION NUMBER: PCT/GB00/04972
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: GB 9930477.6
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Streptomyces sp.
US-10-168-663-7

```

```

RESULT 4
US-10-017-471A-16
; Sequence 16, Application US/10017471A
; Publication No. US20030124644A1
; GENERAL INFORMATION:
;
; APPLICANT: Takano, Eriko
; APPLICANT: Bibb, Mervyn
; TITLE OF INVENTION: Antibiotic Production
; FILE REFERENCE: 0380-P02329US1
; CURRENT APPLICATION NUMBER: US/10/017,471A
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,561
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 16
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Streptomyces coelicolor
US-10-017-471A-16

```

	Query Match	40.9%;	Score 476;	DB 14;	Length 215;
	Best Local Similarity	47.3%;	Pred. No. 3.4e-41;		
	Matches	97;	Conservative 36;	Mismatches 72;	Indels 0; Gaps 0;
QY	11	ROBRAVRTQAI	VRAAASVFDEYGFPAATVAEILSRASVTKGAMYHFHFAKSEELARGVLA	70	
DB	3	KQDRAITPTQTIL	DLAAQVFEKQGYQAAITETLKVAGVTKGALYHFHFSKEELALGVFD	62	
QY	71	EQTLHVAVPESGSKA	QELVDLTMVLVAGHMLHDPILRAGTFLALDQGNVDFSDANPFGEWG	130	
DB	63	AOEPPQAVPEQPL	UQELIDMGMLFCHRLTNTNVAVRAGVRLSMDQQAHGHDRLRGPFFRRWH	122	
QY	131	DICAAQLLAAOER	GEVLPHPVNPKKTGDFIVCGCTGTAQAVSRVTSDRQDLGHRISVMNHHV	190	
DB	123	ETLLKLLNQAKENG	ELLPHVTTDSADLYVGTAGIQVTSQTVSDYQDLHRYVALLQKHI	182	
QY	191	LPSIVPASMLTW	ITETGEERIGKVA	215	
DB	183	LPAIAVPFSLAAL	DLDLSEERGLAA	207	

Query Match	40.4%;	Score	470.5;	DB	15;	Length	276;
Best Local Similarity	45.4%;	Pred. No.	1.8e-40;				
Matches	104;	Conservative	36;	Mismatches	78;	Indels	11;
Gaps	3						
Qy	11	ROERAVRTQATVRAAASVFDEYGEATVAEILSRASVTKGMVYFHPASKEELARGVLA	70				
Db	3	KQRAVQTWRSIVDRAASVDFDYGERRAISEILRRRAKVTGALYFHPASKEATAQAIMD	62				
Qy	71	EQTLHVAVPESGSKAQLVDLTMLVAHGMLHDPILRAGTRTLALDQGAVFDSANPFGEWG	130				
Db	63	EQTSTVEPQEGSPQLSLVDGGQFPAPALRINSNARAGTRLSI--AGVFLGGPHPWGDWI	126				
Qy	131	DICAOILLAAQERGEVLPVHVPKPTGDFIVGCFGTGLQASVRSVTSDDRLGHRISVMKNHV	190				
Db	121	DATARMLELQBERGEVFPQIDPMWSAKIVASFTGILQVSEADSGRADLREQVAMWRHI	180				
Qy	191	LSPTVPASMLTWIETGBERIGKVAATAAEAE-----AAEASEASD	231				
Db	181	LPESIAHPGVIAHIKP-EGRVDLAAQAREKAREEQEARIAAEAKGASD	228				

```

RESULT 7
US-10-156-761-11242
; Sequence 11242, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11242
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11242

```

```

US-10-156-761-11242
Query Match      37.8%; Score 439.5; DB 14; Length 234;
Best Local Similarity 42.6%; Pred. No. 2.5e-37;
Matches 95; Conservative 44; Mismatches 77; Indels 7; Gaps 4

Qy 11 ROERAVRTRQAIIVRAAASVFDEYGEFAATVAEILSRASVTGKAMYHFASKEELARGVLA 70
Db 3 ROERAIRTRQITLVAIAAEVFDEYGEYAATISDVLKRSVGTGKALYFHTSQBELAQAVLA 62
Qy 71 EOTLHV-AVPSGSKAQSLVLTMLVAHGLMH---DPILRAGTRLALDQCAV--DFSDAN 124
Db 63 EQVASLPRVPEQLKQSLDEALLAHLEEGTGDPIVQGSVRLTVDQGSPRDLNRRV 122
Qy 125 PFGEWGDITCAOLLABAQERGEVILPHNPKTKGDFIVCGCTGLQAVRSVTSRDROLGHRIS 184
Db 123 PMQAWTEHTQSLFEERAKGEILPHADVEALAKLFGAFTGCVQLSRIMTGRADLAERVA 182
Qy 185 VMNHNVLPSIVPASMLTWIEGTERRIGKV-AAAAEAAEAAEAS 226
Db 183 DLYRHLMSFPAMGILVRLDFSPGRGSRVYEAMKQRESAAAS 225

RESULT 8
US-10-168-663-5

```

```
; Sequence 5, Application US/10168663
; Publication No. US20040086962A1
; GENERAL INFORMATION:
; APPLICANT: Plant Bioscience Limited
; APPLICANT: Chater, Keith F
; APPLICANT: Bruton, Celia J
; APPLICANT: O'Rourke, Sean J
; APPLICANT: Wietzorrek, Andreas W
; TITLE OF INVENTION: Methods and Materials Relating to Gene Expression
; FILE REFERENCE: 0380-P02909US0
; CURRENT APPLICATION NUMBER: US/10/168,663
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: PCT/GB00/04972
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: GB 9930477.6
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Streptomyces sp.
US-10-168-663-5

Query Match      23.4%; Score 272; DB 15; Length 215;
Best Local Similarity 35.1%; Pred. No. 6.7e-20;
Matches 68; Conservative 35; Mismatches 85; Indels 6; Gaps 2;

QY 11 ROERAVRTRQAIIVRAAASVDFDEYGEFAATVAEILSRASVTKGAMYFHFASKEELARGVLA 70
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3 RQLRAEQTRATIIIGAAADLFDPRGYESTTISEIVAHAGVTGKALYFHFPAKEDLAHAILE 62
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 71 EQ---TLHVAVPESG---SKAQELVDLTMLVAHGMHLDPILRAGTRLALDQGVDFSDAN 124
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 63 IQSRTRSRLAKDLDRGYSSLEALMRLTFGMARLCVQGVPLRAGRLATAGVVPVPPPLPH 122
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 125 PFGEWGDI CAQLLAQAQERGEVLPHVNPCKTGDTFVGCFTGLQAVSRVTSRQDLGHRIS 184
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 123 PFTDWREIATSRLLDAVRQSDVHQDIDVDSVAHTLVCSVVGTVVGGTLEPAGRPRLA 182
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 185 VMNHNVLPSIVPAS 198
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 183 EMWYILIRGMVPT 196
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
US-10-168-663-6
; Sequence 6, Application US/10168663
; Publication No. US20040086962A1
; GENERAL INFORMATION:
; APPLICANT: Plant Bioscience Limited
; APPLICANT: Chater, Keith F
; APPLICANT: Bruton, Celia J
; APPLICANT: O'Rourke, Sean J
; APPLICANT: Wietzorrek, Andreas W
; TITLE OF INVENTION: Methods and Materials Relating to Gene Expression
; FILE REFERENCE: 0380-P02909US0
; CURRENT APPLICATION NUMBER: US/10/168,663
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: PCT/GB00/04972
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: GB 9930477.6
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Streptomyces sp.
US-10-168-663-6

Query Match      22.1%; Score 257; DB 15; Length 215;
Best Local Similarity 34.0%; Pred. No. 2.5e-18;
```

```
Matches 66; Conservative 35; Mismatches 87; Indels 6; Gaps 2;

QY 11 ROERAVRTRQAIIVRAAASVDFDEYGEFAATVAEILSRASVTKGAMYFHFASKEELARGVLA 70
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3 RQLRAEQTRATIIIGAAADLFDPRGYESTTISEIVAHAGVTGKALYFHFPAKEDLAHAILE 62
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 71 EQ---TLHVAVPESG---SKAQELVDLTMLVAHGMHLDPILRAGTRLALDQGVDFSDAN 124
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 63 IQSRTRSRLAKDLDRGYSSLEALMRLTFGMARLCVQGVPLRAGRLATAGVVPVPPPLPH 122
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 125 PFGEWGDI CAQLLAQAQERGEVLPHVNPCKTGDTFVGCFTGLQAVSRVTSRQDLGHRIS 184
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 123 PFTDWREIATSRLLDAVRQSDVHQDIDVDSVAHTLVCSVVGTVVGGTLEPAGRPRLA 182
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 185 VMNHNVLPSIVPAS 198
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 183 EMWYILIRGMVPT 196
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
US-10-156-761-11239
; Sequence 11239, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11239
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11239

Query Match      21.3%; Score 248.5; DB 14; Length 218;
Best Local Similarity 33.3%; Pred. No. 1.9e-17;
Matches 68; Conservative 35; Mismatches 84; Indels 17; Gaps 5;

QY 11 ROERAVRTRQAIIVRAAASVDFDEYGEFAATVAEILSRASVTKGAMYFHFASKEELAR---- 66
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3 KQERAAARTHEALIRSAHAFAERQGYTOARLADISACAGVSPGALHFFHESKAEVARAVEA 62
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 67 --GVLAEQTLHVAVPESGSKAQELVDLTMLVAHGMHLDPILRAGTRLALDQ---GAVDFS 121
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 63 AAGVSLRRAAWLAQPPGTNALQRLTNTSHALERLGDVVARAGFRLNCESAGGGAL--- 119
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 122 DANPFGEWGDI CAQLLAQAQERGEVLPHVNPCKTGDTFVGCFTGLQAVSRVTSRQDL--G 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 120 --NLLREWQTCVEQLLAEEGLIARLRVADTVASVVAATGTFELLGR--RDPFWLSG 175
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 HRISVMNHNVLPSIVPASMLTWIE 204
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 176 QSLAAAFWRVLLPRAATAAALTAVD 199
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
US-10-168-663-11
; Sequence 11, Application US/10168663
; Publication No. US20040086962A1
; GENERAL INFORMATION:
; APPLICANT: Plant Bioscience Limited
```



```

; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9807
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9807

```

```

Query Match      18.5%; Score 215.5; DB 14; Length 218;
Best Local Similarity 29.1%; Pred. No. 5.4e-14;
Matches 57; Conservative 39; Mismatches 85; Indels 15; Gaps 3;

QY 10 VRQERAVRTRQAIIVRAAASVDFEYGFEEATVAEILSRASVTKGAMYFFHFASKEELARGVL 69
Db 27 LKQQRAIRTRGILNRAAFAFATDGPQVTIKDIADGAEMTKGAVYFFHFPSKEALAVL 86
QY 70 ABQTLHVAVPESGSKAQE-----LVDLTMLVAHGMLHDPILRAGTRLALDQAVDF 120
Db 87 EE--FYRMOEAVNGALEHGDPSTPTVVDMRLARAFHEDVFIHAGARLQIERPYIKA 144
QY 121 SDANPFGWGDICQALLAAQERGEVLPVHNPKKTGDFIVGCTGLQAVSRVTSRDQLG 180
Db 145 ELPVPYVGTGLVLTLLDQCRTAGNLPKSTKPEALARALGSVFGAQHISWVINDREDIV 204
QY 181 HRISVMNHNHVLPSIVP 196
Db 205 ERV-----EEIDAFVP 216

```

```

RESULT 15
US-10-156-761-11240
; Sequence 11240, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11240
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11240

```

```

Query Match      17.8%; Score 207.5; DB 14; Length 287;
Best Local Similarity 24.6%; Pred. No. 5.4e-13;
Matches 70; Conservative 36; Mismatches 70; Indels 109; Gaps 8;

QY 10 VRQERAVRTRQAIIVRAAASVDFEYGFEEATVAEILSRASVTKGAMYFFHFASKEELA---- 65
Db 5 VRQERAIRTNALIESAALFAFHGVEVYVLSLTSISARAGVNGALHFFHFPSKAALAEAVG 64
QY 66 -----RGVLAEQ-----TLHVAVP 79

```

Search completed: March 6, 2005, 22:37:43
Job time : 130 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2005, 22:12:12 ; Search time 41 Seconds
(without alignments)
544.446 Million cell updates/sec

Title: US-10-049-710A-2
Perfect score: 1164
Sequence: 1 MAVRHERVAVRQERAVRTRQ.....VAAABAAZAAZAAZAAASDE 232

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1164	100.0	232	2 A57507	virginiae butanol
2	363.5	31.2	226	2 T44586	gamma-butyrolacton
3	267	22.9	215	2 T42057	cprA protein - Str
4	257	22.1	215	2 T35863	A-factor receptor
5	220.5	18.9	213	2 T44588	probable transcrip
6	169.5	14.6	201	2 AG2376	transcription regu
7	138	11.9	196	2 F90919	repressor for uid
8	138	11.9	196	2 C85768	glucuronide repres
9	138	11.9	196	2 D64918	transcription regu
10	135.5	11.6	203	2 C83773	transcription regu
11	135	11.6	189	2 H96503	probable transcrip
12	135	11.6	213	2 H98332	probable transcrip
13	134	11.5	197	2 B83485	probable transcrip
14	134	11.5	207	2 AC2950	transcription regu
15	133	11.4	190	2 S38906	hypothetical prote
16	132.5	11.4	225	2 C70500	probable transcrip
17	132.5	11.4	546	2 G84076	NADH dehydrogenase
18	130.5	11.2	209	2 T34970	probable transcrip
19	128.5	11.0	265	2 A87345	transcription regu
20	128	11.0	198	2 E69779	transcription regu
21	127.5	11.0	208	2 E70947	probable helix-tur
22	127.5	11.0	210	2 D87172	transcription regu
23	126.5	10.9	182	2 AD3214	transcription regu
24	126.5	10.9	247	2 AD3649	transcription regu
25	126	10.8	189	2 D70848	probable transcrip
26	126	10.8	215	2 A87340	transcription regu
27	124.5	10.7	202	2 B87340	transcription regu
28	124.5	10.7	220	2 D85991	probable transcrip
29	124.5	10.7	220	2 B65119	potential acrf/en

30	124.5	10.7	220	2 H91145	probable transcrip
31	123.5	10.6	179	2 B70391	transcription regu
32	123.5	10.6	216	2 C70649	hypothetical prote
33	122	10.5	229	2 F75482	transcription regu
34	122	10.5	295	2 AE1398	transcription regu
35	121	10.4	192	2 C70487	transcription regu
36	121	10.4	199	2 D98212	probable tetr-fam
37	121	10.4	207	2 T36937	probable transcrip
38	120.5	10.4	202	2 AH1382	transcription regu
39	120.5	10.4	203	2 T36403	probable tetr-fam
40	120.5	10.4	209	2 AI3455	transcription regu
41	119.5	10.3	227	2 S27686	transcription regu
42	119	10.2	195	2 AE3074	transcription regu
43	119	10.2	230	2 A95886	probable transcrip
44	118.5	10.2	203	2 S74474	hypothetical prote
45	118	10.1	280	2 C70617	probable transcrip

ALIGNMENTS

RESULT 1

A57507
virginiae butanolide-binding protein - Streptomyces virginiae
C:Species: Streptomyces virginiae
C>Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 09-Jul-2004
C:Accession: A57507
R:Okamoto, S.; Nakamura, K.; Nihira, T.; Yamada, Y.
J. Biol. Chem. 270, 12319-12326, 1995
A>Title: Virginiae butanolide binding protein from Streptomyces virginiae. Evidence tha
A:Reference number: A57507; MUID:95263588; PMID:7744885
A:Accession: A57507
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-232 <OKA>
A:Cross-references: UNIPROT:Q60011; GB:D32251; NID:g961502; PIDN:BAA06981.1; PID:d10075
C:Genetics:
A:Gene: bara

Query Match

Best Local Similarity 100.0%; Score 1164; DB 2; Length 232;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAVRHERVAVRQERAVRTRQAIIVRAAASVFDEYGFEEATVAEILSRASVTKGAMYHFHFS	60
DB	1	MAVRHERVAVRQERAVRTRQAIIVRAAASVFDEYGFEEATVAEILSRASVTKGAMYHFHFS	60
QY	61	KEELARGVLAEOQLHVAVPESGSKAQELVDLTMLVAHGMLHDPILRAGTFLALDQGAADF	120
DB	61	KEELARGVLAEOQLHVAVPESGSKAQELVDLTMLVAHGMLHDPILRAGTFLALDQGAADF	120
QY	121	SDANPFGEGWDICACQLLAEQAERGEVLPVHNPVKTGDFIVGCTGTGLQAVSRVTSRQDLG	180
DB	121	SDANPFGEGWDICACQLLAEQAERGEVLPVHNPVKTGDFIVGCTGTGLQAVSRVTSRQDLG	180
QY	181	HRISVMNHNVLPSIVPASMLTWIETGEERIGKVAIAAAAEAAAEAAEASDE	232
DB	181	HRISVMNHNVLPSIVPASMLTWIETGEERIGKVAIAAAAEAAAEAAEASDE	232

RESULT 2

T44586
gamma-butyrolactone receptor protein tyLP [imported] - Streptomyces fradiae
C:Species: Streptomyces fradiae
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T44586
R:Bate, N.; Butler, A.R.; Gandeche, A.R.; Cundliffe, E.
Chem. Biol. 6, 617-624, 1999
A>Title: Multiple regulatory genes in the tylosin-biosynthetic cluster of Streptomyces
A:Reference number: Z22801; MUID:99398833; PMID:10467127
A:Accession: T44586
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

Db 119 DWAPRWQWQSWVEALRRIERDGLLAGVSAADAATAVAVTAGFEVLG-ENERWLSEE 177
Qy 182 RISSVMNHVLPSPVAPASMLTIWTGTEERIGKVGAAAAAEAA 223
Db 178 RVTGFWNLLPRLTEGRV-----PRARPAGAAASEPAAPA 212

RESULT 6
AG2376
transcription regulator alr4567 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AG2376
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2376
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-201 <KUR>
A:Cross-references: UNIPROT:Q8YNJ8; GB:BA000019; PIDN:BAB76266.1; PID:gl7133703; GSPDB:G
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr4567

Query Match 14.6%; Score 169.5; DB 2; Length 201;
Best Local Similarity 25.5%; Pred. No. 1e-06;
Matches 49; Conservative 41; Mismatches 85; Indels 17; Gaps 5;

Qy 18 TROAIVRAAASVFDEYGFEAATVAEILSRASVTKGAMYPHFASKEELARG---VLAEO 73
Db 7 TKRILQQAELFNQGYAGSSISDINRVTGLQKGGIYNHFQSKDELALQAFYAIASIS 66
Qy 74 LHVAVP-ESGSKAQELVDLTMLVAHGLMLHPILRAGTRLALDQGA VDFSDANP----- 125
Db 67 KHYRIALSRKNAIERVQALIAVFFSPAENPPPIPGCGPLL--NTAVESDDAHALRERTQ 124
Qy 126 --FGWGDICALLAEQAERGEVLPHVNPKTGDFIVGCTGLQAVSRVTSRDQLGHRI 183
Db 125 QAMNSWLNMIRRIQTGIERGEIPEHPYNAEEIATIIVATLEGAIMMSQLYDDTIYMQRAV 184
Qy 184 SVMNHNVLPSIV 195
Db 185 NHU-NHYTENIL 195

RESULT 7
F90919
repressor for uid operon [imported] - Escherichia coli (strain O157:H7, substrain RIMD O
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: F90919
R:Hayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: F90919
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-196 <HAY>
A:Cross-references: UNIPROT:Q59431; GB:BA000007; PIDN:BA835749.1; PID:gl33361793; GSPDB:G
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECe2326

Query Match 11.9%; Score 138; DB 2; Length 196;
Best Local Similarity 26.3%; Pred. No. 0.00037;
Matches 51; Conservative 31; Mismatches 72; Indels 40; Gaps 8;

QY	12	QERAVRTRQALIVRAAAAVFDEYGPBAATVAEILTSRASTVKGAMYFHFASKBELARG-VLA	70
Db	6	QTEAQPTTRILNAAAREIFSENGFHSAMKAICKSCAISPGLTYHHFISKEALIQAILIQ	65
QY	71	BQTLHVA---VPESG-----SKAQELVDLT-----MLVAHGMLDHPILRAGTRL	111
Db	66	DQERALARFRPIEGIHFDVYMSIVSLTHFAFCQRALIVVEIMAEQM-RNQVAAMLK-	123
QY	112	ALDQCAVDSDANPPGEGWDICQAQLABAAQERGEVLPHVNPKKTG-----DFIVGCFTG	167
Db	124	-----NKHMTITFEFVAQMRDAQKQGEISPDIINTAMTSLRLDLLDTGVGLADIE	171
QY	168	A--VSRVTISDRQDL	179
Db	172	AEDLAREASFAQGL	185

RESULT 8

repressor for uid operon [imported] - Escherichia coli (strain O157:H7, substra

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C:Accession: C85768

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D. iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dinalanta, E.; Potamousis, K. Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: C85768

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-196 <STO>

A:Cross-references: UNIPROT:Q59431; GB:AE005174; NID:g12515604; PIDN:AAG56607.

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: uidR

```

Query Match      11.9%; Score 138; DB 2; Length 196;
Best Local Similarity 26.3%; Pred. NO. 0.00037;
Matches 51; Conservative 31; Mismatches 72; Indels 40; Gaps 8

Qy 12 QERAVTRQALVRAAASVFDYGFEEATVAEILSPASVTKGAMYFHPFASKEELARG-VLA 70
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6 QTEAQPTRTRILNAAEIFSENGFHSASMKAIKSCAISPGTLYHHFISKEALIQAIIQ 65
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 71 EOTLHVA--VPESG----SKAEIVDLT-----MLVAHGLMHDPIILAGTRL 111
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66 DQERALARFREPIGIIHFVDMVESIVSLTAEAGQALVVEIMAEGM-RNQVAAMLK- 123
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 112 ALDOGAVDFSDANPFGWGDICAOQLLAERAGGEVLPHVNPKKTG----DFTVGCFGTGLQ 170
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 -----NKHMTITEFVAQMRMDAQKGELSPDINTAMTSRLLLDLTYGVLADE 171
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 168 A--VSRVTSRQDL 179
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 172 AEDLAREASFAQGL 185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
D64918
glucuronide repressor gusR - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C:Accession: D64918
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: D64918
A:Status: nucleic acid
A:Molecule type: DNA
A:Molecule type: translation not shown; translation not shown

```

A:Residues: 1-196 <BLAT>
A:Cross-references: UNIPROT:Q59431; GB:AE000257; GB:U00096; NID:g1787898; PIDN:AACT74690.
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: gusR; uidR
C:Function:
A:Description: repressor of uidRABC (= gusRABC) operon
C:Keywords: transcription regulation

Query Match 11.9%; Score 138; DB 2; Length 196;
Best Local Similarity 26.3%; Pred. No. 0.00037;
Matches 51; Conservative 31; Mismatches 72; Indels 40; Gaps 8;

QY 12 QBRVTRQAIIVRAAASVDEYGFEEATVAEILSRASVTGKAMVFHFASKEELARG-VIA 70
Db 6 QTEAQTRILNAREIPSENGFHSASKKCAICKSCAISPGLYHHFISKEALIQAILQ 65

QY 71 EQLHVA---VPESG-----SKAQLVDLT-----MLVAHGMHLDPILRAGTRL 111
Db 66 DQERALARPREPIEGTHFVDYVVESIVSLTHEAFGORALVVEIMAEGM-RNPQVAAMLK- 123

QY 112 ALDQGAVDPSDANPFGEMGIDICQALLAEQAQERGEVLPHVNPCKTG----DFIVGCTGLQ 167
Db 124 -----NKMWTTEFVAQRMDRAQQRGEISPDINTANTSRLLDLTYGVLAIE 171

QY 168 A--VSRVTSRDQDL 179
Db 172 AEDLAREASFAOGL 185

RESULT 10
C83773
Transcription regulator (Tetr/AcrR family) BH0987 [imported] - Bacillus halodurans (strain C83773)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: C83773
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and A:Reference number: A83650, MUID:20512582; PMID:11058132
A:Accession: C83773
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-203 <STO>
A:Cross-references: UNIPROT:Q9KE70; GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BAB047
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH0987

Query Match 11.6%; Score 135.5; DB 2; Length 203;
Best Local Similarity 28.5%; Pred. No. 0.00062;
Matches 51; Conservative 22; Mismatches 55; Indels 51; Gaps 7;

QY 7 RVAVQER-AVTRQAIIVRAAASVDEYGFEEATVAEILSRASVTGKAMVFHFASKEE--- 63
Db 3 RTKEQNERMSMATREKIYKAALSFLFAKGFALTITKIDISREAHISTGLVYRHFHSKEELF 62

QY 64 --LARGVLAEQTLHVAVPESGSKAQELVD-----LTMVAHGMHLDP 103
Db 63 GLRIEKAIGEMSVTKFLETDTSPKELVSEFVTMTAGIQSSEVSHYFLLVARSLLEDE 122

QY 104 IL-RAG-----TRLALDQGAVDPSDANPFGEMGIDICQALLAEQAQERGEVLPHVNPCKTG 157
Db 123 VLPKIGDLKTKTDIALFQQ-----MAELIARGQRQGEF-----KEGD 158

RESULT 11
H96903
Transcription regulator Tetr/AcrR family [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: H96903
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C1C
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: H96903
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-189 <KUR>
A:Cross-references: UNIPROT:Q97N07; GB:AE001437; PIDN:AAK78019.1; PID:g15022852; GSPDB:C
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC0032

Query Match 11.6%; Score 135; DB 2; Length 189;
Best Local Similarity 26.7%; Pred. No. 0.00062;
Matches 46; Conservative 31; Mismatches 53; Indels 42; Gaps 7;

QY 17 RTRQAIIVRAAASVDEYGFEEATVAEILSRASVTGKAMVFHFASKEELARGVLAQTLHV 76
Db 3 KTKNAIFKSALMIFSKNGYDGAITWDEIASNAKVAKTLYHYHFKSKEEIFKYVISE-GMNV 61

QY 77 AVPSGSKAQELVDLTMLVAHGMHLDP1--LRAGTRLALDQGAV---DFSDANPFGBWG 130
Db 62 IREMEQEA-----GKSNFVNKLKAIICRFQI--GMIFNRDFFPKVLSQLWG 107

QY 131 -----DICAQLLAEQAQERGEVLPHVNPCKTGDFIVGCTGL 166
Db 108 QDSRQLELRKVKIDYIEIIRYLVKDAEMKGYI-----KKGETEFMSYTFPGV 154

RESULT 12
H98332
Probable transcription regulator PA3574 [imported] - Agrobacterium tumefaciens (strain C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: H98332
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B. Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: H98332
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-213 <KUR>
A:Cross-references: UNIPROT:Q8UB16; GB:AE007870; PIDN:AAK90186.1; PID:g15160193; GSPDB:C
C:Genetics:
A:Gene: AGR_L 3219
A:Map position: linear chromosome
C:Superfamily: probable transcription repressor mtrr

Query Match 11.6%; Score 135; DB 2; Length 213;
Best Local Similarity 31.0%; Pred. No. 0.00072;
Matches 44; Conservative 28; Mismatches 38; Indels 32; Gaps 6;

QY 8 VAVQER-AVTRQAIIVRAAASVDEYGFEEATVAEILSRASVTGKAMVFHFASKEELA 65
Db 5 ISMRRTKAAEETQAIILAAAEAVFVKGVANSLSDEVAAGVTRGAIYWHFSSKADLF 64

QY 66 RGLVLAEQTLHVAVPESGSKAQELVDLTMLVAHGMHLDPILRAGTRLALDQGAVDPSDANP 125
Db 65 IGLYES----VALPES-----DLID-----FGDPLKGTALLAKIEAT----- 99

QY 126 FGEWGDICQALLAEQAQERGEVL 147
Db 100 -CKW---LALLAEDEQRQIRIM 116

RESULT 13
B83485
Probable transcription regulator PA1290 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

This Page Blank (uspo)


```
RESULT 2
O24741
ID O24741 PRELIMINARY; PRT; 221 AA.
AC O24741;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Para.
GN Name=fara;
OS Streptomyces sp.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1931;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRI-5;
RX MEDLINE=97405912; PubMed=9260956;
RA Waki M., Nihira T., Yamada Y.;
RT "Cloning and characterization of the gene (fara) encoding the receptor
RT for an extracellular regulatory factor (IM-2) from Streptomyces sp.
RT strain FRI-5.";
RL J. Bacteriol. 179:5131-5137(1997).
CC -!- SIMILARITY: Contains 1 HTH tetR-type DNA-binding domain.
DR EMBL; AB001683; BA021859.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS01081; HTH TETR_1; 1.
KW DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 221 AA; 24282 MW; E928FC54AE23PD7 CRC64;

Query Match 42.5%; Score 495; DB 2; Length 221;
Best Local Similarity 46.8%; Pred. No. 3.4e-32;
Matches 101; Conservative 40; Mismatches 75; Indels 0; Gaps 0;

QY 12 QRAVTRQAIIVRAASVDEYGFEEATVAEILSRASVTGKAMVFHFKASKEELARGVLA 71
Db 4 QVRAIRTRQAILSAARVDEYGFEEATVAEILSRASVTGKALYFHFKASKEELARGVLA 63

QY 72 QTLHVAVPESGSKAQLVDLTMLVAHGMHDPILRAGTRALDQGVAVDFSDANPFGWGD 131
Db 64 QNEDLLPPEPAKQLEVDVAVMLHTRLTNPVAVGVLSLDVAVAGGLDRAPFNWVD 123

QY 132 ICAQLLAQAQERGEVLPHVNPVKTKGDFIVGCTGLQAVSRVTSDDRLGHRISVMNVHL 191
Db 124 KETDLLKQAQAQCELLPHVVPVPAETADVTGAYGVQSMQALTEHQDLGQRVNALRLHLM 183

QY 192 PSIVPASMLTWIETGEERIGKVAAAEAEAEAE 227
Db 184 PSIAQPSVLASLHLSGEAEVLEARQLARQADE 219

RESULT 3
O86852
ID O86852 PRELIMINARY; PRT; 215 AA.
AC O86852;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Gamma-butyrolactone binding protein.
GN Name=scrB; ORFNames=SCAH10.30c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Eubacteria, and M145;
```

```
Takano E., Chakraborty R., Nihira T., Yamada Y., Bibb M.;
Submitted (AUG-1998) to the EMBL/GenBank/DBSJ databases.
RA
RL
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -!- SIMILARITY: Contains 1 HTH tetR-type DNA-binding domain.
DR EMBL; AJ007731; CAA07628.1; -.
DR EMBL; AL939127; CAB60184.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS01081; HTH TETR_1; 1.
KW Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 215 AA; 23861 MW; C5D30C679408ABB9 CRC64;

Query Match 40.9%; Score 476; DB 2; Length 215;
Best Local Similarity 47.3%; Pred. No. 1.1e-30;
Matches 97; Conservative 36; Mismatches 72; Indels 0; Gaps 0;

QY 11 QRAVTRQAIIVRAASVDEYGFEEATVAEILSRASVTGKAMVFHFKASKEELARGVLA 70
Db 3 QKDAIRTRQITLDAQAQVFEKQGYQAATITELKLVAGVTGKALYFHFKASKEELARGV 62

QY 71 EQTLHVAVPESGSKAQLVDLTMLVAHGMHDPILRAGTRALDQGVAVDFSDANPFGW 130
Db 63 AQEPVQAVPEQPLRIQELIDMGMLFCHRLTNVAVAGVRLSMDQQAHLDRRGPFRRWH 122

QY 131 DTCAQLLAQAQERGEVLPHVNPVKTKGDFIVGCTGLQAVSRVTSDDRLGHRISVMNVH 190
Db 123 ETLLKLLNQAKENGELLPHVVTDSADLYVGTGTFAGIQVVSQTVSDYQDLHEHYALLQXHI 182

QY 191 LPSIVPASMLTWIETGEERIGKVAA 215
Db 183 LPAIAVPSVLAAALDLSEERGLAA 207

RESULT 4
AFRP STRGR
ID AFRP STRGR STANDARD; PRT; 276 AA.
AC Q9ZNF78; Q54189;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE A-factor receptor protein (A-factor binding protein).
GN Name=arpa;
OS Streptomyces griseus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1911;
RN [1]
RP SEQUENCE FROM N.A.; SEQUENCE OF 6-21; 44-53 AND 148-167, SUBUNIT.
RP A-FACTOR-BINDING, AND VARIANTS.
RC STRAIN=IFO 13350;
RX MEDLINE=96042082; PubMed=7592371;
RA Onaka H., Ando N., Nihira T., Yamada Y., Beppu T., Horinouchi S.;
RT "Cloning and characterization of the A-factor receptor gene from
```

RT Streptomyces griseus.";

RL J. Bacteriol. 177:6083-6092 (1995).

RN [2]

RP SEQUENCE FROM N.A., DNA-BINDING, A-FACTOR-BINDING, AND MUTAGENESIS OF

RT VAL-41; TRP-119; GLU-135; PRO-138; LYS-147; ILE-149; GLY-155; SER-160;

RP ASP-163; ASP-168 AND PRO-187.

RC STRAIN=JA 5142;

RX MEDLINE=99033010; PubMed=9813285; DOI=10.1016/S0378-1119(98)00487-9;

RA Sugiyama M., Onaka H., Nakagawa T., Horinouchi S.;

RT "Site-directed mutagenesis of the A-factor receptor protein: Val-41

RT important for DNA-binding and Trp-119 important for ligand-binding.";

RL Gene 222:133-144 (1998).

RN [3]

RP SUBCELLULAR LOCATION, AND A-FACTOR-BINDING.

RX MEDLINE=89327144; PubMed=2502536;

RA Miyake K., Horinouchi S., Yoshida M., Chiba N., Mori K., Nogawa N.,

RA Morikawa N., Beppu T.;

RT "Detection and properties of A-factor-binding protein from

RT Streptomyces griseus.";

RL J. Bacteriol. 171:4298-4302 (1989).

RN [4]

RP A-FACTOR-BINDING.

RX MEDLINE=90264288; PubMed=2111804;

RA Miyake K., Kuzuyama T., Horinouchi S., Beppu T.;

RT "The A-factor-binding protein of Streptomyces griseus negatively

RT controls streptomycin production and sporulation.";

RL J. Bacteriol. 172:3003-3008 (1990).

RN [5]

RP DNA-BINDING, AND A-FACTOR-BINDING.

RX MEDLINE=97363713; PubMed=9220006;

RA Onaka H., Horinouchi S.;

RT "DNA-binding activity of the A-factor receptor protein and its

RT recognition DNA sequences.";

RL Mol. Microbiol. 24:991-1000 (1997).

RN [6]

RP FUNCTION.

RX MEDLINE=20010118; PubMed=10540289;

RA Ohnishi Y., Kameyama S., Onaka H., Horinouchi S.;

RT "The A-factor regulatory cascade leading to streptomycin biosynthesis

RT in Streptomyces griseus: identification of a target gene of the A-

RT factor receptor.";

RL Mol. Microbiol. 34:102-111 (1999).

RN [7]

RP MUTANT HOI.

RX MEDLINE=97252506; PubMed=9098075;

RA Onaka H., Sugiyama M., Horinouchi S.;

RT "A mutation at proline-115 in the A-factor receptor protein of

RT Streptomyces griseus abolishes DNA-binding ability but not ligand-

RT binding ability.";

RL J. Bacteriol. 179:2748-2752 (1997).

CC -!- FUNCTION: Represses adpA expression by binding to the promoter

CC region in the absence of A-factor, causing repression of

CC streptomycin production and of sporulation.

CC -!- SUBUNIT: Homodimer or multimer. Binds to both DNA and A-factor as

CC a homodimer.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- DOMAIN: Binds DNA through its N-terminal H-T-H motif and binds A-

CC factor via its C-terminal region.

CC -!- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; D49782; BAA08617.1; --

DR EMBL; AB021882; BAA36282.1; --

DR GO; GO:0005737; C:cytoplasm; IDA.

DR GO; GO:0003700; F:transcription factor activity; IDA.

DR GO; GO:0042174; P:negative regulation of sporulation; IC.

DR GO; GO:0045892; P:negative regulation of transcription, DNA-d. . . ; IDA.

DR GO; GO:0019872; P:streptomycin biosynthesis; IC.

DR InterPro; IPR009057; Homeodomain_like.

DR Pfam; PF00440; Tetr.N; 1.

DR PRINTS; PR00455; HTHtetr.

DR PROSITE; PS01081; HTH_TETR_1; FALSE_NEG.

DR PROSITE; PS00977; HTH_TETR_2; 1.

KW Direct protein sequencing; DNA-binding; Repressor; Sporulation;

RT Transcription; Transcription regulation.

FT DOMAIN 8 68

FT DNA BIND 31 50

FT H-T-H motif (Potential).

FT P -> S (in mutant HOI; lacks DNA-binding

FT activity).

FT G -> E (in strain IFO 13350).

FT AATDSG -> PTSEGT (in strain IFO 13350).

FT GA -> VT (in strain IFO 13350).

FT V -> I (in strain IFO 13350).

FT W->A: Loss of DNA-binding activity.

FT W->A: Loss of A-factor-binding activity.

FT E->A: Loss of both DNA-binding and A-

FT Factor-binding activities.

FT P->A: No loss of DNA-binding or A-factor-

FT binding activities.

FT K->A: No loss of DNA-binding or A-factor-

FT binding activities.

FT I->A: Loss of both DNA-binding and A-

FT factor-binding activities.

FT G->A: Loss of both DNA-binding and A-

FT factor-binding activities.

FT S->A: No loss of DNA-binding or A-factor-

FT binding activities.

FT D->A: No loss of DNA-binding or A-factor-

FT binding activities.

FT D->A: Loss of both DNA-binding and A-

FT factor-binding activities.

FT P->A: No loss of DNA-binding or A-factor-

FT binding activities.

FT SQ SEQUENCE 276 AA; 28950 MW; 6940BC3105D35CE0 CRC64;

Query Match 40.4%; Score 470.5; DB 1; Length 276;

Best Local Similarity 45.4%; Pred. No. 4.2e-30;

Matches 104; Conservative 37; Mismatches 77; Indels 11; Gaps 3;

QY 11 ROERAVRTRQATVRAAASVDFYGEFAATVAEILSRASVTGKAMYPHFASKELARGVLA 70

Db 3 KQARAVQVTRSVDAASVDFDYGERAAISILRRKVTGKALYFHFASKAIAQAIMD 62

QY 71 EQLHVAVPESGSKAQLVDLTMLVAHGMHDPILRAGTRALDQGAVDSDANPGEWG 130

Db 63 EGTSTVEFEQSGPLQSLVDGGQQQAFALRHNSMARAGTRLSIE--GVFLGGPHPGDWI 120

QY 131 DICAOALLAAQSRGEVLPHVNPCKTGDFVCGFTGLQAVSRVTSRDQDLGHRISVMNVH 190

Db 121 DATARMLELQGEVGFPOIDPMVSAKIIVASFTGIQLVSEADSGRADLGGVAEMRWRI 180

QY 191 LPSIVPASMLTWTGTEERIGKVAATAAE-----AAEASEASD 231

Db 181 LPSIAHPGVIAHKP-EGRVLDAAQAAREKAREBEQEARIAAEAKGASD 228

RESULT 5

Q84E47

ID Q84E47 PRELIMINARY; PRT; 232 AA.

AC Q84E47;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Butyrolactone autoregulator receptor protein.

GN Name=scaR;

OS Streptomyces clavuligerus.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycineae; Streptomycetaceae; Streptomyces.

```

OX NCBI_TaxID=1901;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15257430;
RA Kim H.S., Lee Y.J., Lee C.K., Choi S.U., Yeo S.H., Hwang Y.I.,
RA Yu T.S., Kinoshita H., Nihira T.;
RT "Cloning and characterization of a gene encoding the gamma-
RT butyrolactone autoregulator receptor from Streptomyces clavuligerus.";
RL Arch. Microbiol. 182:44-50(2004).
CC -!- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
DR EMBL; AB106894; BAC66444.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009050; Globin like.
DR InterPro; IPR001647; HTH Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS01081; HTH TETR 1; 1.
KW Complete proteome, DNA-binding; Receptor; Transcription;
SQ SEQUENCE 232 AA; 25423 MW; 4EAB640486C7B27B CRC64;

Query Match 37.8%; Score 440; DB 2; Length 232;
Best Local Similarity 45.9%; Pred. No. 1e-27;
Matches 94; Conservative 40; Mismatches 65; Indels 6; Gaps 4;

QY 11 ROERAVTRQAIIVRAAASVFDEYGEAATVAEILSRASVTGKAMYFHFASKELARGVLA 70
DB 3 ROERAIRTRQITLVAAAEVDFQVGEAATISDLQSGUTGKALYFHFASKELAQVLA 62
QY 71 EQTLHV-AVPESGSKAQELVDLTMLVAH---GMLHDPILRAGTRLALDOG-AVDFSDAN- 124
DB 63 EQVAALPPVPPQGLKLEAIDEGMLLAHLTGEHGDPIIQSGIRLTVDQSSKNDLDRV 122
QY 125 PFGEWGDICAQLLAAQAERGEVLPVNPVKTTGDFIVGCTGLQAVSRVTSRDQDLGHRIS 184
DB 123 PMQGWIDHSLGLFSEARKNGEVLPHADLDSVSRFLFAGCTGVQVLSRIMTDRQDLAERMS 182
QY 185 VMNHNVLPSIVPASMILTWTETGEER 209
DB 183 DMIRNLMPALAVPSVLRLDFSTDR 207

RESULT 6
Q82H41
ID Q82H41 PRELIMINARY; PRT; 234 AA.
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative gamma-butyrolactone receptor protein.
GN Name=avar; OrderedLocusNames=SAV3705;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=WA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,

RA NCBI_TaxID=1901;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15257430;
RA Kim H.S., Lee Y.J., Lee C.K., Choi S.U., Yeo S.H., Hwang Y.I.,
RA Yu T.S., Kinoshita H., Nihira T.;
RT "Cloning and characterization of a gene encoding the gamma-
RT butyrolactone autoregulator receptor from Streptomyces clavuligerus.";
RL Arch. Microbiol. 182:44-50(2004).
CC -!- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
DR EMBL; AB106894; BAC66444.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009050; Globin like.
DR InterPro; IPR001647; HTH Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS01081; HTH TETR 1; UNKNOWN 1.
KW DNA-binding; Receptor; Transcription; Transcription regulation.
SQ SEQUENCE 232 AA; 25423 MW; 4EAB640486C7B27B CRC64;

Query Match 37.8%; Score 439.5; DB 2; Length 234;
Best Local Similarity 42.6%; Pred. No. 1.2e-27;
Matches 95; Conservative 44; Mismatches 77; Indels 7; Gaps 4;

QY 11 ROERAVTRQAIIVRAAASVFDEYGEAATVAEILSRASVTGKAMYFHFASKELARGVLA 70
DB 3 ROERAIRTRQITLVAAAEVDFQVGEAATISDLVLRSGVTGKALYFHFASKELAQVLA 62
QY 71 EQTLHV-AVPESGSKAQELVDLTMLVAHGMHLH--DPILRAGTRLALDOGAV--DFSDAN 124
DB 63 EQVASLPVRPEQELKQSLDQSLDQSLDQSLDQSLDQSLDQSLDQSLDQSLDQSLDQSLDQSL 122
QY 125 PFGEWGDICAQLLAAQAERGEVLPVNPVKTTGDFIVGCTGLQAVSRVTSRDQDLGHRIS 184
DB 123 PMQWTEHTQSLFEERAKGEILPHADVEALAKLVGAGTGVQVLSRIMTGRADLAERVA 182
QY 185 VMNHNVLPSIVPASMILTWTETGEERIGKV-AAAAEAAEAAEAS 226
DB 183 DLYRLMPSFAMPGLVRLDFSPERSRVYEAAMQRESAAAS 225

RESULT 7
Q83X22
ID Q83X22 PRELIMINARY; PRT; 222 AA.
AC Q83X22;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative Tetr family receptor protein.
OS Streptomyces rochei (Streptomyces parvullus).
OG Plasmid pSLA2-L.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1928;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7434AN4;
RX MEDLINE=22676866; PubMed=12791134;
RA Mochizuki S., Hiratsu K., Suwa M., Ishii T., Sugino F., Yamada K.,
RA Kinashi H.;
RT "The large linear plasmid pSLA2-L of Streptomyces rochei has an
RT unusually condensed gene organization for secondary metabolism.";
RL Mol. Microbiol. 48:1501-1510(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=7434AN4;
RX MEDLINE=20408175; PubMed=10954087;
RA Hiratsu K., Mochizuki S., Kinashi H.;
RT "Cloning and analysis of the replication origin and the telomeres of
RT the large linear plasmid pSLA2-L in Streptomyces rochei.";
RL Mol. Gen. Genet. 263:1015-1021(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=7434AN4;

```

RX MEDLINE=20331737; PubMed=10767533; DOI=10.1016/S0378-1119(00)00060-3;
RA Suwa M., Sugino H., Sasaka A., Mori E., Fujii S., Shinkawa H.,
RA Nimi O., Kinashi H.,
RT "Identification of two polyketide synthase gene clusters on the linear
RT plasmid pSLA2-L in Streptomyces rochei.",
RL Gene 246:123-131(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=7434AN4;
RX MEDLINE=99053144; PubMed=9836424;
RA Kinashi H., Fujii S., Hatanai A., Kurokawa T., Shinkawa H.;
RT "Physical mapping of the linear plasmid pSLA2-L and localization of
RT the eryA and actI homologs.",
RL Biotechnol. Biochem. 62:1892-1897(1998).
CC -!- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
DR EMBL; AB088224; BAC76540.1; -.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR001647; HTH Tetr.
DR Pfam; PF00440; Tetr.N.1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS01081; HTH_TETR_1; UNKNOWN 1.
DR DNA-binding; Plasmid; Receptor; Transcription;
KW Transcription regulation.
SQ SEQUENCE 222 AA; 23805 MW; EBE16185AFB85F09 CRC64;

Query Match 37.3%; Score 434.5; DB 2; Length 222;
Best Local Similarity 42.7%; Pred. No. 2.8e-27;
Matches 94; Conservative 45; Mismatches 78; Indels 3; Gaps 3;

QY 11 ROERAVRTRQAIIVRAAASVDFEYGEFAATVAEILSRASVTGKAMYPHFASKELARGVLA 70
Db 3 QQERAIRTRRAVLEAAATVFAEYGAATVADILKVGATKGLYFHPFSKEALRGILE 62

QY 71 EQTLHVAVPESGSKAQELVDLTMLVAHGMHLDPILRAGTRRLALDQGVDFSDANPFGWG 130
Db 63 AQVPPQLVPPQ-LKIQEWDAGMTLAHQLPDRPVPVAGARLSAETHGSE-QHGSAPPTWI 120

QY 131 DICAOILAAOERGEVLPVHNPKTGDFIVGCTGLQAVSRVTSRDQDLGHRISVNMVHV 190
Db 121 AFSASLLEQKNGEVGLGIEFAETACVIGSGFHGILQLLSQLTQNWADIEQASALFRHV 180

QY 191 LPSIVPASMLTWTETGEERIGKVAASAAEAAEASAAAS 230
Db 181 LPAAVPSVLVRLDTAPDRGARVVALEAM-AAQPDGLAS 219

RESULT 8
Q6VMG9
ID Q6VMG9 PRELIMINARY; PRT; 237 AA.
AC Q6VMG9;
DT 05-JUL-2004 (TEMBLrel. 27, Created)
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE Putative gamma-butyrolactone binding protein.
GN Name=alpZ;
OS Streptomyces ambofaciens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1889;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23877;
RX PubMed=14742212;
RA Pang X., Aigle B., Girardet J.M., Mangenot S., Pernodet J.L.,
RA Decaxis B., Leblond P.;
RT "Functional anquacycline-like antibiotic gene cluster in the terminal
RT inverted repeats of the Streptomyces ambofaciens linear chromosome.",
RL Antimicrob. Agents Chemother. 48:575-588(2004).
CC -!- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
DR EMBL; AY338477; AAR30170.1; -.

DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR001647; HTH Tetr.
DR Pfam; PF00440; Tetr.N.1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS01081; HTH_TETR_1; 1.
DR DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 237 AA; 25701 MW; F649FA2A64AE5834 CRC64;

Query Match 36.8%; Score 428; DB 2; Length 237;
Best Local Similarity 40.5%; Pred. No. 1e-26;
Matches 90; Conservative 47; Mismatches 85; Indels 0; Gaps 0;

QY 11 ROERAVRTRQAIIVRAAASVDFEYGEFAATVAEILSRASVTGKAMYPHFASKELARGVLA 70
Db 3 QQERAIRTQLSILKAAAEVDFSHGYEATIGEILRAGVTGKALYFHPFSKQALAEVLE 62

QY 71 EQTLHVAVPESGSKAQELVDLTMLVAHGMHLDPILRAGTRRLALDQGVDFSDANPFGWG 130
Db 63 QQFSVIRVPPGFCRLQEFVDTGLIVAYRMRRDPLVSAVARLSLEQELRAEYSGSAIRQWI 122

QY 131 DICAOILAAOERGEVLPVHNPKTGDFIVGCTGLQAVSRVTSRDQDLGHRISVNMVHV 190
Db 123 GASEVLGLAAKEQGLLPHVPAESAWLFSAAWTGTQLYSQILGLEDLEERVVALFRHL 182

QY 191 LPSIVPASMLTWTETGEERIGKVAASAAEAAEASAAASDE 232
Db 183 LPSIAPVAVLSLEITERRAARLGSASDALLGPGEQADSPQ 224

RESULT 9
Q8VVP1
ID Q8VVP1 PRELIMINARY; PRT; 228 AA.
AC Q8VVP1;
DT 01-WAR-2002 (TEMBLrel. 20, Created)
DT 01-WAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-WAR-2004 (TEMBLrel. 26, Last annotation update)
DE Autoregulator receptor protein.
GN Name=spBR;
OS Streptomyces pristinaespiralis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=38300;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21570198; PubMed=11557748; DOI=10.1074/jbc.M101109200;
RA Folcher M., Gaillard H., Nguyen L.T., Nguyen K.T., Lacroix P.,
RA Bamas-Jacques N., Rinkel M., Thompson C.J.;
RT "Pleiotropic functions of a Streptomyces pristinaespiralis
RT autoregulator receptor in development, antibiotic biosynthesis, and
RT expression of a superoxide dismutase.",
RL J. Biol. Chem. 276:44297-44306(2001).
CC -!- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
DR EMBL; AY026762; AAK07686.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR Pfam; PF00440; Tetr.N.1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS01081; HTH_TETR_1; UNKNOWN 1.
DR DNA-binding; Receptor; Transcription; Transcription regulation.
KW SEQUENCE 228 AA; 25012 MW; 1CC9EF160C756CB5 CRC64;

Query Match 35.7%; Score 415; DB 2; Length 228;
Best Local Similarity 43.9%; Pred. No. 1.1e-25;
Matches 98; Conservative 39; Mismatches 76; Indels 10; Gaps 6;

QY 11 ROERAVRTRQAIIVRAAASVDFEYGEFAATVAEILSRASVTGKAMYPHFASKELARGVLA 70
Db 3 QQERAVRTRRAILVAAAEVDFSHGYEATIGEILKESGLTKGALYFHPFSKEELAQVLA 62

QY 71 EQTLHVA--VPESGSKAQELVDLTMLVAHGMHLH---DPILRAGTRRLALDQGVDFSDAN 124

```
Db 63 EQ-VHALPDLPEGLMLQTAVDRAILLALHLLRRDTGDPVIRGSRVLTVEQALRDGLDR 121
Qy 125 -PFGWGDICAOILAERGEREVLPHVNPCKTGDFIVGCTGLQAVSRVTSRDOLGHR 183
Db 122 VPMQAMEGTQDLFOQAAGSILPHVDIVGAAKTFVGAGTGVQLSNIMTGRQDMTERV 181
Qy 184 SYMNHVLPISVPASMLTWIETGEERIGKVAASAAAEAAEAS 226
Db 182 ADLYRFLMTAIVPGVLVLDSPGR--GVLAEEAVRRDAA 222

RESULT 10
Q9RPK9 PRELIMINARY; PRT; 220 AA.
AC Q9RPK9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE TarA.
OS Streptomyces tendae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC31160;
RA Engel P., Scharfenstein L.L.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
DR EMBL; AF156161; AAF06961.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR InterPro; IPR009057; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS01081; HTH_TETR_1; UNKNOWN 1.
KW DNA-binding; Receptor; Transcription; Transcription regulation.
SQ SEQUENCE 220 AA; 24010 MW; DCEABE24C6581D02 CRC64;
```

```
Query Match 34.7%; Score 403.5; DB 2; Length 220;
Best Local Similarity 44.0%; Pred. No. 9e-25;
Matches 92; Conservative 41; Mismatches 71; Indels 5; Gaps 4;

Qy 11 ROERAVRTRQAIIVRAAASVFDEYGFEAATVAEILSRASVTKGAMYPHFASKKEELARGVLA 70
Db 3 QODRAVRTRAVIRAAAVFAERGAAATISILKRGVTKGALYFHFDSKAALAQGV 62

Qy 71 EQ-TLHVAVPESGSKAQELVDLTMLVAHGMLHDPILRAGTRIALDQ-GAVDFSDANPFGE 128
Db 63 EQLTPEYHLPRE-LKLQEWVDAGMTLARRLPREPFLLAGVRISADRPGRGVLGSAWP--A 119

Qy 129 WGDICAOILAERGEREVLPHVNPCKTGDFIVGCTGLQAVSRVTSRDOLGHRISVMWN 188
Db 120 WSLRSHVLTAKRGREVLPHVVPPEETAQVFLGAVVGAQFVSQTLAGWEDLDRTAALYS 179

Qy 189 HVLPSIVPASMLTWIETGEERIGKVAASAA 217
Db 180 HLLGAIAAPVILRLDTAPDGRARVIAE 208
```

```
RESULT 11
Q84H96 PRELIMINARY; PRT; 208 AA.
AC Q84H96;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Butyrolactone receptor.
OS Streptomyces carzinostaticus subsp. neocarzinostaticus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
```

```
OX NCBI_TaxID=167636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15944;
RA Liu W., Nonaka K., Nie L., Bae J., Zazopoulos E., Parnet C.M.,
RA Shen B.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
DR EMBL; AY117439; AAM78022.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS01081; HTH_TETR_1; UNKNOWN 1.
KW DNA-binding; Receptor; Transcription; Transcription regulation.
SQ SEQUENCE 208 AA; 23076 MW; F328B30B25E7507B CRC64;
```

```
Query Match 34.5%; Score 402; DB 2; Length 208;
Best Local Similarity 42.5%; Pred. No. 1.1e-24;
Matches 79; Conservative 37; Mismatches 70; Indels 0; Gaps 0;

Qy 11 ROERAVRTRQAIIVRAAASVFDEYGFEAATVAEILSRASVTKGAMYPHFASKKEELARGVLA 70
Db 5 QODRAVRTRALILRAAAEVFDESGYSGASISKIMNRAGVTQCGMYFHFPSKKRLALAVNS 64

Qy 71 EQTLHVAVPESGSKAQELVDLTMLVAHGMLHDPILRAGTRIALDQ-GAVDFSDANPFGEWG 130
Db 65 SQOTFIDFPFSEGAGLQVRVIDLTFLHAELOTNPLIRASIRLAVEGEGFQVRDDTAYQDWW 124

Qy 131 DICACQALLAERGEREVLPHVNPCKTGDFIVGCTGLQAVSRVTSRDOLGHRISVMWNHV 190
Db 125 ALVETYLLEARQRGELLPEVDEHDFAVVLVSFTGSQSFSSVASARADLPISRIVLLWRYL 184

Qy 191 LPSIVP 196
Db 185 LPGLTP 190
```

```
RESULT 12
Q6X4A0 PRELIMINARY; PRT; 220 AA.
AC Q6X4A0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative gamma-butyrolactone-binding protein.
OS Streptomyces ansochromogenes.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=115647;
RN [1]
RP SEQUENCE FROM N.A.
RA Li W., Tan H.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 HTH tetr-type DNA-binding domain.
DR EMBL; AY256849; AAP87374.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS01081; HTH_TETR_1; 1.
KW DNA-binding; Receptor; Transcription; Transcription regulation.
SQ SEQUENCE 220 AA; 23920 MW; B76FF63127595602 CRC64;
```

```
Query Match 34.5%; Score 401; DB 2; Length 220;
Best Local Similarity 42.5%; Pred. No. 1.4e-24;
Matches 94; Conservative 40; Mismatches 79; Indels 8; Gaps 4;
```



```
QY 11 RQRAVTRQAIIVRAAASVDFEYGFPAATVAEILSPASVTGKAMYFHPASKKEELARGVLA 70
DB 3 QQDRAVTRERAVLRAAFAAFAERGAAYATSEILKKAAGVTGKALYFHPDSKAALAQGVLIQ 62
QY 71 EQ-TLHVAVPESGSKAQELVDLTMLVAHGMHLDPILRAGTRALALDQAVDFSDANPFGEW 129
DB 63 EQLTPEYHLPRE-LKLOEWVDAGMALARLPRPFFLLAGVRISADRPGRDVL-GSANPAW 120
QY 130 GIDCAQLLAAQERGEVLPHPNPKTKGDFIVGFTGLQAVSRVTSRDQLGHRISVMVNH 189
DB 121 ARLTSHALTEAKRGEVLPHPVPEETAQVFLGAWGAFVQSOTLAGWADLDDRTAALYGH 180
QY 190 VLPSIVPASMLTWIETGEERIGKVAASAAEAAEASAAAS 230
DB 181 LLAAIAAPPVLRLDTPDRGARV-----IAEARRESDLS 216

RESULT 13
Q6REPF8
ID Q6REPF8 PRELIMINARY; PRT; 221 AA.
AC Q6REPF8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PDK3.089;
OS Rhodococcus sp. DK17.
OG Plasmid pDK3.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Rhodococcus.
OX NCBI_TaxID=186196;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DK17;
RA Choi K.Y., Kim D., Sul W.J., Chae J.-C., Zylstra G.J., Kim Y.M.,
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 HTH tetra-type DNA-binding domain.
DR EMBL; AY502076; AAR90230.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTEPR.
DR DNA-binding; Hypothetical.
KW Transcription regulation.
SQ SEQUENCE 221 AA; 24204 MW; E52906C946B94CAE CRC64;

Query Match 33.5%; Score 390; DB 2; Length 221;
Best Local Similarity 40.5%; Pred. No. 1.le-23;
Matches 79; Conservative 39; Mismatches 71; Indels 6; Gaps 1;

QY 10 VRQRAVTRQAIIVRAAASVDFEYGFPAATVAEILSPASVTGKAMYFHPASKKEELARGV 69
DB 2 VQQRRAVTRQAIIVRAAEMFDSKGFEGASLVEILEAGTKGALYFHPFSKEDLARHII 61
QY 70 AEQ-----TLHVAVPESGSKAQELVDLTMLVAHGMHLDPILRAGTRALALDQAVDFSDA 123
DB 62 AEQHRISISAVQAIAQAESAIEQIVMLCHEMARQIVQDPPIVRAGIRITLSEADDRGPA 121
QY 124 NPFGEGDICAQLLAAQERGEVLPHPNPKTKGDFIVGFTGLQAVSRVTSRDQLGHRIRI 183
DB 122 GPYLDIAACQTLARAVAEGDLVDITDPTFARYVIGATGVTQVTSQVLTHTRADLEQRV 181
QY 184 SVMNHNVLPSIVPAS 198
DB 182 DENWGFLLPGIMETT 196

RESULT 14
Q6L812
ID Q6L812 PRELIMINARY; PRT; 217 AA.
```

```
AC Q6L812;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE KsBA (Fragment).
GN Name=kSBA;
OS Kitasatospora setae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Kitasatospora.
OX NCBI_TaxID=2066;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 14216; and IFO14216;
RX PubMed=15150228; DOI=10.1128/JB.186.11.3423-3430.2004;
RA Choi S., Lee C., Hwang Y., Kinoshita H., Nihira T.;
RT "Cloning and functional analysis by gene disruption of a gene encoding
a gamma-butyrolactone autoregulator receptor from Kitasatospora
setae.";
RL J. Bacteriol. 186:3423-3430(2004).
CC -!- SIMILARITY: Contains 1 HTH tetra-type DNA-binding domain.
DR EMBL; AB126048; BAD20239.1; -.
DR EMBL; AB121071; BAD20239.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTEPR.
DR DNA-binding; Transcription.
KW DNA-binding; Transcription regulation.
SQ SEQUENCE 217 AA; 23755 MW; 54C7579CCF7C33DA CRC64;

Query Match 32.1%; Score 374; DB 2; Length 217;
Best Local Similarity 35.2%; Pred. No. 2.2e-22;
Matches 77; Conservative 51; Mismatches 81; Indels 10; Gaps 2;

QY 14 RAVTRQAIIVRAAASVDFEYGFPAATVAEILSPASVTGKAMYFHPASKKEELARGVLA 73
DB 6 RAAKTREALIIQAAGVFEQHGNGARLSMITSRSLGTWGAIVFHTSKDDLARAVNAAQA 65
QY 74 LHVAVPESGSKAQELVDLTMLVAHGMHLDPILRAGTRALALDQAVDFSDANPFGEWGDIC 133
DB 66 DDLVLPDGERGLQRLDITLYLASELQRLNVLRLAGVRLAEQGSFGVRDSTPYLQWVDRF 125
QY 134 AQLLAAQERGEVLPHPNPKTKGDFIVGFTGLQAVSRVTSRDQLGHRISVMNHNVLPS 193
DB 126 RDELGAEEAVGOLLPEVVVADVVAQLLVSSYQTLLSEIATDRADLPERIVRMWTVLLPG 185
QY 194 IVPASMLTWIETGEERIGKVAASAAEAAEASAAASDE 232
DB 186 IA-----TPEART-RLRLEPSVADRADDRGNDQ 214

RESULT 15
Q9XCX7
ID Q9XCX7 PRELIMINARY; PRT; 226 AA.
AC Q9XCX7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Gamma-butyrolactone receptor protein TyLP.
GN Name=tyLP;
OS Streptomyces fradiae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1906;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T59235;
RX MEDLINE=993998833; PubMed=10467127; DOI=10.1016/S1074-5521(99)80113-6;
RA Bate N., Butler A.R., Gandeche A.R., Cundliffe E.;
RT "Multiple regulatory genes in the tylosin biosynthetic cluster of
Streptomyces fradiae.";
RL Chem. Biol. 6:617-624(1999).
```

